Asset Return Correlations in Episodes of Systemic Crises

by

David Melkuev

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Quantitative Finance

Waterloo, Ontario, Canada, 2014

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis explores asset return correlation dynamics in relation to systemic crises. The eigenvalues obtained from principal component analysis performed on the sample return correlation matrix equal the variance explained by their associated eigenvectors. Also referred to as the absorption ratio in extant literature, the proportion of total system variance captured by a relatively small number of principal components reflects the strength of co-movement in the system. As a stylized fact of financial returns, increasing correlation is known to coincide with or follow significant market drawdowns. However, studies in recent years have suggested that contagion is facilitated by a state of high interconnectedness which precedes systemic crises.

We study the association between changes in the absorption ratio and aggregate market returns in the context of three crisis episodes. Time series of the normalized eigenvalue estimates reveal that the structure of return correlations is richly dynamic across various asset groups. Using linear regression models, we find that systemic crises are characterized by a general breakdown of the correlation structure rather than increased co-movement. In addition, we find that changes in the normalized eigenvalues of the correlation matrix do not have consistent predictive power for market returns or volatility. Evidence for the relationship is stronger on a contemporaneous and lagging basis. Particularly, changes in correlation are significantly related to realized returns and volatility in the same period, for different period lengths. In general, realized volatility has a more robust relationship with shifts in the correlation structure than returns.

We also use a nonparametric technique to monitor divergence in the distributions underlying successive observations of the normalized dominant eigenvalue. Periods of high divergence imply a change in the correlation structure and are found to either precede or coincide with systemic shocks.

Acknowledgements

I would like to thank my supervisor, Professor Tony Wirjanto, for his advice and accommodative guidance throughout my graduate studies and specifically this thesis. Under his supervision I was able to explore a broad range of topics and follow my interests to more focused research. Professor Wirjanto was always available for discussions, through which he helped improve my work and revealed an inspiring breadth of knowledge. I am also thankful to my readers, Professors Adam Kolkiewicz and Chengguo Weng, for providing valuable feedback. Professor Song Liu made freely available software which was useful for Chapter 4 and this is thankfully acknowledged. I am grateful to Alonso Gomez, who has provided valuable perspectives and support for this thesis as well as career matters. Financial support from the BMO Capital Markets Lime Connect partnership is gratefully acknowledged.

The huge amount of gratitude I owe to my family and friends for support on this thesis is only a drop in the ocean compared to everything they have given me.

Any errors are my own.

Table of Contents

List of Tables

List of Figures

List of Symbols

X, x	Scalar-valued random variable, sample
$\boldsymbol{X}, \boldsymbol{x}$	Vector-valued random variable, sample
$\mathcal{X}, \mathbf{X}$	Matrix-valued random variable, sample
$\mathbb{E}[X]$	Expected value operator
$\mathrm{Var}[X]$	Variance operator
$\mathrm{Cov}[\cdot, \cdot]$	Covariance operator
$\mathrm{Cor}[\cdot, \cdot]$	Pearson product-moment correlation operator
$(\cdot)'$	Matrix transpose operator
$P(\cdot)$	Probability of an event
$\xrightarrow{\mathfrak{D}}$	Convergence in distribution
$\mathbf{I}_d$	$d \times d$ identity matrix
$\mathbf{1}_n'$	$n \times 1$ vector of ones
$\mathcal{O}(d)$	Group of orthogonal $d \times d$ matrices
$\mathbb{R}^{d \times n}$	Family of real $d \times n$ matrices
$\det(\cdot)$	Determinant of a matrix
$\mathrm{diag}(\cdot)$	Diagonal matrix
$\mathrm{etr}(\cdot)$	Exponential trace of a matrix
$\mathrm{tr}(\cdot)$	Trace of a matrix
$\{\cdot\}_T$	Time series with index set T
Γ_d	d-variate gamma distribution

$N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	d-variate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$
$W_d(n, \boldsymbol{\Sigma})$	Wishart distribution with n degrees of freedom and covariance matrix $\boldsymbol{\Sigma}$

Chapter 1

Introduction

Much of the observable phenomena in the empirical sciences are of multivariate nature. Similarly in finance, values of multiple assets are observed simultaneously and form a system with complex dynamics. A good understanding of valuations of multiple assets jointly is critical to a wide range of financial applications in areas such as portfolio optimization, risk management and asset pricing. For instance, since pioneering work in portfolio management, the notion of risk in finance has been defined by variance of asset returns and the extension to the multivariate setting of covariances is natural due to the co-movement that asset values exhibit. This co-movement is often referred to as correlation, although correlation describes a linear relationship and, most of the time, assuming it to be such is an oversimplification of the true relationship between returns on different assets[1]. Nevertheless, statistical correlation sheds some light on features of financial markets that are fundamental and universal, and is therefore an important topic of research. As an example at a basic level, that the values of different assets within the equity class are generally positively correlated points to the presence of class-specific risk factors. In the remainder of this thesis, unless otherwise stated, we will use the terms *correlation* and *co-movement* interchangeably to refer to the extent that observable quantities fluctuate together.

In this thesis we will explore the phenomenon of changes in the strength of co-movement between asset returns. What can be perhaps most reliably said about this is that correlations tend to increase during downturns. There is plenty of empirical evidence supporting this assertion, for example in [6, 15, 19, 51, 55, 77] and the references therein. According to some of these studies, the prevalence of asymmetric correlations with respect to direction of returns warrants that it be regarded as a stylized fact in stock (and possibly other) markets. Stylized facts are properties of financial markets that are so robust – for example across a wide range of economies, instruments and time periods – that they are accepted as generally true to the extent that some market models tend to emulate them. Through an extensive inquiry into asymmetric correlation, it is most commonly believed to be a co-incident effect, the reasons for which tend

[1] The most common use of correlation in statistics is in reference to the Pearson product-moment correlation coefficient.

to be behavioural, like increased linkages due to forced liquidations under leverage. Indeed correlation asymmetry has been linked to asymmetric volatility, a stylized fact also known as the leverage effect. Such conclusions are robust to market fluctuations of varying amplitude, commensurately with the definition of stylized facts. In contrast, we study return correlations in reference to more circumstantial market drawdowns. Specifically, we are interested in correlation dynamics related to systemic shocks.

Systemic events have a special characteristic of causing drawdowns that are contagious in that they ultimately propagate to a large number of assets relative to the number of assets immediately affected. In other words, under certain market conditions devaluation of some assets can cause the devaluation of many other assets. The notion of financial contagion is not new and is recognized in historic accounts of financial crises dating back several centuries [70]. Considering systemic crises alone, it is intuitive to explain correlation asymmetry as a manifestation of financial contagion itself. But the risk of systemic shocks changes over time as the conditions that facilitate contagion are created. Therefore, a central question posed in this thesis is whether increased systemic risk is reflected in rising correlations before the onset of a systemic crisis.

To measure the strength of co-movement of asset returns, we take an approach based on spectral analysis. The idea is to transform the coordinates of the original multivariate return data such that the variance along each axis of the new system is maximized. The values along each axis represent linearly uncorrelated factors, and have successively decreasing variance. This transformation is called principal component analysis (PCA) and is a popular technique in multivariate statistical analysis. In particular, it has also been used to gain insight into changes in correlation between asset returns [14, 20, 29, 43, 54, 60, 90]. The vectors of values along each axis are called principal components (PCs). The orthogonal linear combinations for the required transformation are found as the eigenvectors of the system's correlation matrix, each of which has an associated eigenvalue. Since each eigenvalue equals the variance of the PC found through its associated eigenvector, we can define the total risk in the system as the sum of these eigenvalues. Then, the proportion of total variance attributed to a small number of PCs indicates the degree of commonality between returns on the assets in question. We estimate the proportion of variance explained by a fixed small number of factors through time and study its fluctuations. A similar statistic is named the absorption ratio (AR) in [43], in reference to the proportion of variance absorbed by a few factors. Studies that have taken this approach commonly find that severe downturns coincide with and are sometimes preceded by increases in the strength of co-movements. The explanation provided for this is that stronger integration allows for the ripple effect that characterizes systemic financial crises. Literature drawing such conclusions is based on experiments that are limited to certain asset classes over certain time periods. Specifically, the Financial Crisis of 2008 is by far the most common case study for this framework and equities are the most common asset class examined. This thesis will expand the analysis by examining three separate crises of systemic nature: the Financial Crisis of 2008, the Eurozone Sovereign Debt Crisis and the Asian Financial Crisis of 1997. Furthermore, a greater range of assets will

be investigated in this thesis, including equities, bonds, credit default swaps and currencies.

In addition, it is important to recognize that sample eigenvalues should be treated as random variables and we thus discuss distributional aspects of eigenvalue-based statistics. Results from multivariate statistical analysis include exact expressions about both joint and marginal eigenvalue distributions of Wishart matrices (see [57], for examples.) Although Wishart matrices are useful models of correlation matrices, there are issues in practical applications of exact density representations. Firstly, the true covariance matrix is a required parameter that is usually not known. Secondly, for a non-null covariance structure the expressions are difficult to evaluate, especially in high-dimensional settings. Yet results from random matrix theory (RMT) point to a convergence of the density of sample eigenvalues to a non-random density as matrix dimensions go to infinity at a fixed aspect ratio. The cornerstone result for Wishart matrices has been known since 1967 as the Marčenko-Pastur law [53]. Since then, the limiting eigenvalue density has been shown to exist for many different matrix structures, including covariance matrices of time series with temporal dependence [39, 89] or cross sectional dependence [75]. Knowledge that the limiting eigenvalue density exists implies that covariance stationarity in a series of eigenvalue observations can be a reasonable assumption under some conditions and this is a desirable property for statistical inference. Furthermore, the discrepancy between empirical and theoretical eigenvalues points to a deviation of the actual system covariance structure from what is postulated in theory. For example, if the empirical largest eigenvalue exceeds the upper support of the Marčenko-Pastur density then this serves as evidence that there is more structure to the data than in multivariate white-noise.

Widespread crises can cause significant socioeconomic damage and policy makers therefore have a strong interest in detecting elevation of systemic risk as soon as possible in order to take preventive action. Some market-based indicators have exhibited potential to act as early warning signals [37]. If changes in correlation in asset returns have such potential then it is of interest to test the lead-lag relationship between changes in correlation and asset returns. This issue is addressed in [14, 54] but no formal tests between the eigenvalue magnitude and returns are made. Instead, it is noted qualitatively that the proportion of variance explained by the first PC increased before the crisis studied. We fill this gap as well using statistical tests based on linear models.

Finally, recognizing drawbacks of the linearity assumption for returns, we also explore a nonparametric method to gain insight into change-points in the correlation of returns. The problem of detecting changes in the underlying distribution of a stochastic process is commonly known as *change-point detection* and is a classical one. Its origin can be traced back as early as the 1930s in work on the problem of monitoring the quality of manufacturing processes [84]. More recently, this problem has been studied in a wide array of fields including econometrics [3, 4, 5, 12, 16] and finance [10, 18, 73]. There are many detection procedures and perhaps at the highest level, they differ in being *offline* or *online*. Offline algorithms are algorithm and are applied on a sample of historical observations. The purpose of the offline algorithms is to detect changes

in an *a posteriori* fashion within the sample with the lowest probability of error. On the other hand, online algorithms are designed for real-time monitoring for changes in a stochastic process. Online detection procedures aim to minimize both the false alarm rate and the detection delay, as each new sample is assumed to incur a cost. We will take an online approach to monitoring the strength of asset co-movements. However, instead of searching for an algorithm with a decision function on whether change has occurred, we propose to track a measure of divergence between the distributions of successive observations. This addresses robustness issues that often arise in financial application due to financial data being inherently noisy.

1.1 Contribution

We conduct a novel empirical investigation of the connection between changes in correlation and systemic risk in three cases of financial crises: the Financial Crisis of 2008, the Eurozone Sovereign Debt Crisis and the Asian Financial Crisis of 1997. To measure correlation we adapt a variant of the AR, which is computed from the eigenvalues of the sample return correlation matrix. Such spectral analysis of correlation has been previously used in studies of equities during the Financial Crisis. We expand on this by investigating different episodes of crisis as well as a broader set of assets. We find that the relationship between correlation and extreme drawdowns is more general than what has been shown in extant literature. Namely, whereas extreme drawdowns have been linked to increasing co-movement in asset returns, we find that in certain cases a systemic shock results in a decoupling of asset returns. Thus systemic crises are associated with a general breakdown of correlation structure.

We conduct formal statistical tests about the linear relationship between changes in correlation and financial crises. We find that there is no sufficient evidence of a robust lead-lag effect between changes in correlation and returns based on linear regression and Granger-causality tests. Changes in the AR do not have predictive power for the realized return or volatility of a broad market index. However, realized volatility is consistently and significantly related to changes in the AR across different time frequencies and asset groups. We statistically confirm that the directional relationship can be both negative and positive.

In light of the evident non-stationarity in correlation structure we conduct change-point analysis using nonparametric technique. By estimating the divergence in the distributions of successive groups of samples through time, we identify time periods that are associated with shifts in the correlation structure. We find that the divergence score increases either before or coincidentally with systemic financial shocks.

1.2 Content Structure

We structure this thesis in four chapters as follows. In the remainder of this chapter we present basic mathematical notions and definitions that will be used extensively throughout this thesis. In Chapter 2 we present a spectral approach to measuring the strength of co-variation in multivariate series. In this chapter we also discuss some statistical properties of this measure. Chapter 3 contains an empirical study of the correlation dynamics in reference to three financial crises of systemic nature. The relationship between changes in correlations of various assets and aggregate market returns are investigated. In Chapter 4 we take a nonparametric approach based on statistical change-point detection techniques to study the non-stationarity of correlations. We end with concluding remarks.

1.3 Basic Definitions

We state here basic notions and definitions relating to main concepts used throughout this thesis.

1.3.1 Covariance Matrices

Definition 1 (Covariance matrix)**.** Let $\boldsymbol{X} = [X_1 \ldots, X_d]'$ and $\boldsymbol{Y} = [Y_1 \ldots, Y_d]'$ be d-dimensional vector-valued random variables with mean vectors $\boldsymbol{\mu}_X$ and $\boldsymbol{\mu}_Y$. The *covariance matrix* of $\boldsymbol{X}$ and $\boldsymbol{Y}$ is defined to be a $d \times d$ symmetric matrix

$$\boldsymbol{\Sigma}_{XY} \equiv \mathrm{Cov}\big[\boldsymbol{X}, \boldsymbol{Y}\big] = \mathbb{E}\left[(\boldsymbol{X} - \boldsymbol{\mu}_X)(\boldsymbol{Y} - \boldsymbol{\mu}_Y)'\right] \tag{1.1}$$

and its (i, j)-th entry $\sigma_{X_i Y_j}$ equals the covariance between X_i and Y_j.

More commonly we are interested in the covariance of $\boldsymbol{X}$ with itself,

$$\boldsymbol{\Sigma}_X \equiv \mathrm{Cov}\big[\boldsymbol{X}, \boldsymbol{X}\big] = \mathbb{E}\left[(\boldsymbol{X} - \boldsymbol{\mu}_X)(\boldsymbol{X} - \boldsymbol{\mu}_X)'\right] \tag{1.2}$$

which is a $d \times d$ symmetric covariance matrix of $\boldsymbol{X}$ with the (i, j)-th entry equaling the covariance between variates X_i and X_j whereas the i-th diagonal equals the variance of X_i. This is a multivariate analogue of the variance concept in one dimensional random variables, so we represent their equivalence by the notation $\mathrm{Var}\,[\boldsymbol{X}] \equiv \mathrm{Cov}\,[\boldsymbol{X}, \boldsymbol{X}]$.

The covariance structure of a multivariate random variable is generally invariant under translations for distributions with existing and finite second moments [11]. That is, for constant $d \times 1$

vectors $\boldsymbol{\delta} = [\delta, \ldots, \delta]'$ and $\boldsymbol{\eta} = [\eta, \ldots, \eta]'$, the covariance matrix of $\boldsymbol{U} = \boldsymbol{X} + \boldsymbol{\delta}$ and $\boldsymbol{V} = \boldsymbol{Y} + \boldsymbol{\eta}$ is

$$\begin{aligned}
\boldsymbol{\Sigma}_{UV} &= \mathbb{E}\left[(\boldsymbol{X} + \boldsymbol{\delta} - \boldsymbol{\mu}_{X+\delta})\left(\boldsymbol{Y} + \boldsymbol{\eta} - \boldsymbol{\mu}_{Y+\eta}\right)'\right] && (1.3)\\
&= \mathbb{E}\left[(\boldsymbol{X} + \boldsymbol{\delta} - \boldsymbol{\mu}_X - \boldsymbol{\delta})\left(\boldsymbol{Y} + \boldsymbol{\eta} - \boldsymbol{\mu}_Y - \boldsymbol{\eta}\right)'\right] && (1.4)\\
&= \mathbb{E}\left[(\boldsymbol{X} - \boldsymbol{\mu}_X)(\boldsymbol{Y} - \boldsymbol{\mu}_Y)'\right] && (1.5)\\
&= \boldsymbol{\Sigma}_{XY}. && (1.6)
\end{aligned}$$

Therefore, we may assume without loss of generality that $\boldsymbol{\mu}_X = \boldsymbol{\mu}_Y = \mathbf{0}$ in equation (1.1).

Definition 2 (Pearson correlation coefficient)**.** Let X and Y be random variables with respective variances σ_X and σ_Y. The *Pearson correlation coefficient* between X and Y is

$$\rho_{XY} \equiv \operatorname{Cor}[X, Y] = \frac{\operatorname{Cov}[X, Y]}{\sigma_X \sigma_Y} \tag{1.7}$$

We will frequently refer to the multivariate extension of the notion of (Pearson) correlation between the variates of a random vector, which is matrix-valued.

Definition 3 (Correlation matrix)**.** The *correlation matrix* of a random vector $\boldsymbol{X} = [X_1, \ldots, X_d]$ is a $d \times d$ symmetric matrix whose (i, j)-th element equals $\rho_{X_i X_j}$. In matrix form,

$$\mathbf{P}_X = \boldsymbol{\sigma}^{-1} \boldsymbol{\Sigma}_X \boldsymbol{\sigma}^{-1}, \tag{1.8}$$

where $\boldsymbol{\sigma}$ is a $d \times d$ diagonal matrix with σ_{X_i} in the i-th diagonal element.

In most practical settings, population parameters such as means or covariance matrices are unknown and are estimated using available observations of the random variable of interest. Thus all definitions above have sample counterparts.

Definition 4 (Sample covariance matrix)**.** Given a sample of n observations $\mathbf{X} = \begin{bmatrix} \boldsymbol{x}_1 & \ldots & \boldsymbol{x}_n \end{bmatrix}$, the *sample covariance matrix* is

$$\mathbf{S}_X = \frac{1}{n-1}\left(\mathbf{X} - \bar{\mathbf{X}}\mathbf{1}_n'\right)\left(\mathbf{X} - \bar{\mathbf{X}}\mathbf{1}_n'\right)' \tag{1.9}$$

where $\bar{\mathbf{X}} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} \boldsymbol{x}_i = \frac{1}{n}\mathbf{X}\mathbf{1}_n$.

The sample covariance matrix is an unbiased estimator of the population covariance matrix. Replacing $\boldsymbol{\Sigma}_X$ with $\mathbf{S}_X$ in equation (1.8), we obtain the sample correlation matrix, denoted by $\mathbf{R}_X$. Importantly, if the observations are standardized to have unit variance, then $\mathbf{S}_X = \mathbf{R}_X$.

Given a linear combination $\boldsymbol{c}'\boldsymbol{X}$ of the variates of $\boldsymbol{X}$, we have

$$\begin{aligned} \text{Var}[\boldsymbol{c}'\boldsymbol{X}] &= \boldsymbol{c}'\text{Var}[\boldsymbol{X}]\boldsymbol{c} && (1.10)\\ &= \boldsymbol{c}'\boldsymbol{\Sigma}\boldsymbol{c} && (1.11)\\ &\geq 0 && (1.12) \end{aligned}$$

where the inequality is true because variance is non-negative. This means that valid covariance matrices must be positive semi-definite, a characterization defined in the following.

Definition 5 (Positive semi-definite matrix)**.** A real $d \times d$ matrix $\boldsymbol{\Sigma}$ is called *positive semi-definite* if, for every $d \times 1$ vector $\boldsymbol{c} \in \mathbb{R}^d$,

$$\boldsymbol{c}'\boldsymbol{\Sigma}\boldsymbol{c} \geq 0. \tag{1.13}$$

Furthermore if none of the elements in $\boldsymbol{X}$ is a linear combination of the others then the inequality (1.12) holds strictly if and only if $\boldsymbol{c} = \boldsymbol{0}$. We then say that $\boldsymbol{\Sigma}$ is a positive-definite matrix.

Definition 6 (Positive-definite matrix)**.** A real $d \times d$ matrix $\boldsymbol{\Sigma}$ is called *positive-definite* if, for every non-zero $d \times 1$ vector $\boldsymbol{c} \in \mathbb{R}^d$,

$$\boldsymbol{c}'\boldsymbol{\Sigma}\boldsymbol{c} > 0. \tag{1.14}$$

A positive-definite covariance matrix is invertible; however, a covariance matrix that is positive semi-definite but not positive-definite is not invertible.

We will also make extensive use of the eigendecomposition of a square matrix in terms of its pairs of eigenvectors and associated eigenvalues, as defined below.

Definition 7 (Eigendecomposition)**.** The *eigendecomposition* of a square matrix $\boldsymbol{\Sigma}$ in terms of its d eigenvectors, $\boldsymbol{v}_i$, and corresponding eigenvalues, λ_i, is given by

$$\begin{aligned} \boldsymbol{\Sigma} &= \sum_{i=1}^{d} \lambda_i \boldsymbol{v}_i \boldsymbol{v}_i' && (1.15)\\ &= \boldsymbol{\Upsilon}\boldsymbol{\Lambda}\boldsymbol{\Upsilon}', && (1.16) \end{aligned}$$

where

$$\boldsymbol{\Upsilon} = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_d \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_d \end{bmatrix}. \tag{1.17}$$

All eigenvalues of a square symmetric, positive-definite matrix are positive. Hence, by the definitions above, the eigenvalues of a valid invertible covariance matrix are all positive.

1.3.2 Time Series

A multivariate time series model is a stochastic process $\{\boldsymbol{X}_t\}_{t\in\mathbb{Z}}$, i.e. a series of random vectors indexed by the integers and defined on some probability space $(\Phi, \mathcal{F}, P)$. Most time series processes studied in the literature are assumed to be, loosely speaking, constant through time in terms of their distributional properties. That is, if any realization of the process was divided into different intervals, then the behaviour of the process within each interval would be statistically similar. This property is referred to as stationarity and can be formally described in two different senses.

Definition 8 (Strict stationarity)**.** The multivariate time series $\{\boldsymbol{X}_t\}_{t\in\mathbb{Z}}$ is said to be *strictly stationary* if

$$F_{\boldsymbol{X}_s,\ldots,\boldsymbol{X}_t}(\boldsymbol{x}_s,\ldots,\boldsymbol{x}_t) = F_{\boldsymbol{X}_{s+k},\ldots,\boldsymbol{X}_{t+k}}(\boldsymbol{x}_s,\ldots,\boldsymbol{x}_t) \quad \forall s,t,k\in\mathbb{Z}, \tag{1.18}$$

where F is the joint distribution of its suffixes.

Strict stationarity implies that the probabilistic structure is invariant under shifts in time and observation lengths. Requiring this property to hold may be too forbidding from an applied point of view. One example is white noise but there is a dearth of strictly stationary processes in popular domains of application. A weaker requirement of stationarity is typically adopted and it can be stated as follows.

Definition 9 (Weak stationarity)**.** The multivariate time series $\{\boldsymbol{X}_t\}_{t\in\mathbb{Z}}$ is said to be *weakly stationary* if its first two moments exist and satisfy

$$\mathbb{E}\left[\boldsymbol{X}_t\right] = \mathbb{E}\left[\boldsymbol{X}_{t+k}\right] = \boldsymbol{\mu} \qquad \forall t,k\in\mathbb{Z} \tag{1.19}$$

$$\mathrm{Cov}\left[\boldsymbol{X}_s,\boldsymbol{X}_t\right] = \mathrm{Cov}\left[\boldsymbol{X}_{s+k},\boldsymbol{X}_{t+k}\right] \qquad \forall s,t,k\in\mathbb{Z}. \tag{1.20}$$

The definition of weak stationarity implies that the mean vector of the series is constant, while the covariance depends only on the temporal separation between observations. Therefore, $\mathrm{Cov}\left[\boldsymbol{X}_s,\boldsymbol{X}_t\right] = \mathrm{Cov}\left[\boldsymbol{X}_0,\boldsymbol{X}_{t-s}\right]$, where $\tau = t-s$ is called the *lag*, so that a matrix-valued covariance function of the lag argument will suffice in order to fully describe the covariance structure.

Chapter 2

Spectral Analysis of Correlation

This chapter will outline the concept of correlation structure of a multivariate time series that will be referred to throughout the paper. Specifically, we aim to investigate the strength of contemporaneous co-movement between components in the series.

2.1 Eigenspectrum Analysis

Consider a d-dimensional random vector $\boldsymbol{X}$, with

$$\boldsymbol{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_d \end{bmatrix} \sim F \tag{2.1}$$

for some distribution F and linearly independent $X_i, i = 1, \ldots, d$. Let $\boldsymbol{\Sigma} \in \mathbb{R}^{d \times d}$ be the population covariance matrix of $\boldsymbol{X}$. Since $\boldsymbol{\Sigma}$ is a square symmetric matrix, it can be diagonalized to obtain its eigendecomposition

$$\boldsymbol{\Sigma} = \sum_{i=1}^{d} \lambda_i \boldsymbol{v}_i \boldsymbol{v}_i' \tag{2.2}$$

$$= \boldsymbol{\Upsilon} \boldsymbol{\Lambda} \boldsymbol{\Upsilon}'. \tag{2.3}$$

The eigendecomposition of $\boldsymbol{\Sigma}$ is deterministic up to a scaling factor and in reference to any eigenvectors hereafter, we will mean the normalized scale, i.e. $\boldsymbol{v}'\boldsymbol{v} = 1$. It is also unique if $\lambda_i \neq \lambda_j \, \forall i, j$ and we will assume this property to be true.

Computing the eigendecomposition of a covariance matrix amounts to Principal Component Analysis (PCA), which is one of the most frequently used techniques for multivariate statistical analysis. The population PCA is defined as a linear transformation $\boldsymbol{X} \mapsto \boldsymbol{Y}$ such that the components of $\boldsymbol{Y}$, called principal components (PCs), are linearly uncorrelated and have successively

decreasing variance. In particular, the map $\boldsymbol{Y} = \boldsymbol{\Upsilon}'\boldsymbol{X}$ achieves this if we choose to index the eigendecomposition such that $\lambda_1 > \ldots > \lambda_d$. Then, an interpretation of the eigenvectors is that $\boldsymbol{v}_1$ is the direction in $\mathbb{R}^d$ of largest variance in the data; $\boldsymbol{v}_2$ is the direction of largest variance that is orthogonal to $\boldsymbol{v}_1$; and each subsequent $\boldsymbol{v}_i$ points in the direction of greatest variance such that it is orthogonal to the previous eigenvector. This means that the first PC contains the most information about the structure of the system and every subsequent PC contains less information than the previous one.

In addition, if $\boldsymbol{Y}_i = \boldsymbol{v}_i'\boldsymbol{X}$ is the i-th PC, then

$$\begin{aligned}
\mathrm{Var}[\boldsymbol{Y}_i] &= \mathrm{Cov}\big[\boldsymbol{v}_i'\boldsymbol{X}, \boldsymbol{v}_i'\boldsymbol{X}\big] && (2.4)\\
&= \boldsymbol{v}_i'\mathrm{Cov}\big[\boldsymbol{X}, \boldsymbol{X}\big]\boldsymbol{v}_i && (2.5)\\
&= \boldsymbol{v}_i'\boldsymbol{\Sigma}\boldsymbol{v}_i && (2.6)\\
&= \boldsymbol{v}_i'\boldsymbol{v}_i\lambda_i && (2.7)\\
&= \lambda_i, && (2.8)
\end{aligned}$$

so λ_i is the variance of the i-th PC. The eigenvalues of $\boldsymbol{\Sigma}$ are all positive because it is symmetric and positive-definite. Since the principal components are uncorrelated and the variance of each is given by the corresponding eigenvalue, we can define the total variance in the system as

$$\Omega = \sum_{i=1}^{d} \lambda_i, \tag{2.9}$$

and the variance accounted for by the first k PCs as

$$\omega_i = \frac{\lambda_i}{\Omega}. \tag{2.10}$$

Then, the proportion of variance explained by the first k PCs is

$$\phi_k = \sum_{i=1}^{k} \omega_i \qquad k \leq d. \tag{2.11}$$

This framework translates the concentration of cumulative eigenvalue magnitude, or energy, to the strength of co-movement in the components of $\boldsymbol{X}$. The value of ϕ_k can be interpreted as the extent to which the variates in $\boldsymbol{X}$ move together due to some k uncorrelated factors. If a high proportion of total variance is explained by relatively few PCs, then this would manifest in an eigenspectrum with the few corresponding eigenvalues being much larger than the remaining bulk. Large eigenvalues are associated with a structure in the data whereas small eigenvalues are associated with noise. Therefore, highly correlated random variables will result in high values of ϕ_k for a relatively small k.

Consider, for illustration, an extreme case with a bivariate random vector $\boldsymbol{X} = [X_1\ X_2]'$ such

that
$X_1 \sim N(0,1)$ and $X_2 = 2X_1$. The population covariance matrix for this system is

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \tag{2.12}$$

with an eigendecomposition

$$\mathbf{\Upsilon} = \begin{bmatrix} 0.45 & -0.90 \\ 0.90 & 0.45 \end{bmatrix} \quad \text{and} \quad \mathbf{\Lambda} = \begin{bmatrix} 5 & 0 \\ 0 & 0 \end{bmatrix}. \tag{2.13}$$

The first PC is proportional to $[1,2]'$, capturing the full structure of the system, namely variance on the line $X_1 = 2X_2$. The eigenvalues are $\lambda_1 = 5$ and $\lambda_2 = 0$. Therefore for a sample of n observations, ϕ_1 will be close to one. On the other hand, if X_1 and X_2 are independently distributed standard Normal random variables, then $\lambda_1 = \lambda_2 = 1$ and, statistically, the eigenvectors of a sample covariance matrix from this system will explain an approximately equal amount of variance[1].

2.1.1 Sample Eigenspectrum Analysis

In most practical cases the population covariance matrix is unknown so the true eigenvectors and eigenvalues cannot be determined with certainty. Instead, the available information is a sample of n realizations of $\boldsymbol{X}$, which we write in matrix form as

$$\mathbf{X} = \begin{bmatrix} \boldsymbol{x}_1 & \dots & \boldsymbol{x}_n \end{bmatrix} \in \mathbb{R}^{d \times n}. \tag{2.14}$$

The analysis in Section 2.1 has a sample analog. First, $\mathbf{\Sigma}$ is estimated by the sample covariance matrix $\mathbf{S}$, which has an eigendecomposition

$$\mathbf{S} = \sum_{i=1}^{d} l_i \boldsymbol{u}_i \boldsymbol{u}_i' \tag{2.15}$$

$$= \mathbf{U}\mathbf{L}\mathbf{U}', \tag{2.16}$$

where the $\boldsymbol{u}_i$'s are eigenvectors with associated eigenvalues l_i. Unlike (2.3), this decomposition is stochastic and, in fact, $\boldsymbol{u}_i$ and l_i are observation-dependent estimates of $\boldsymbol{\upsilon}_i$ and λ_i, respectively.

[1] It will be shown in Section 2.2 that the eigenvalues are systematically unequal.

Our estimators of (2.9)-(2.11) are then given by

$$\hat{\Omega} = \sum_{i=1}^{d} l_i, \tag{2.17}$$

$$\hat{\omega}_i = \frac{l_i}{\hat{\Omega}}, \tag{2.18}$$

$$\hat{\phi}_k = \sum_{i=1}^{k} \hat{\omega}_i \qquad k \leq d. \tag{2.19}$$

Remark 1. Note that the rows of $\mathbf{X}$ are typically pre-processed to have zero mean. If they were also standardized to have unit variance, then we would have $\mathbf{S} = \mathbf{R}$, where $\mathbf{R}$ is the sample correlation matrix and $\hat{\Omega} = \text{tr}(\mathbf{R}) = d$.

2.1.2 Time Series Eigenspectrum Analysis

Consider a d-dimensional time series $\{\boldsymbol{X}_t\}$ whose population covariance is $\boldsymbol{\Sigma}$. We are interested in performing a sequential analysis of the eigenspectrum of this random process using the framework above. Of particular interest is the quantity in (2.19), which indicates the strength of contemporaneous co-movement between the components of $\{\boldsymbol{X}_t\}$. This can be obtained by applying PCA on a fixed-size rolling sample of observations from $\{\boldsymbol{X}\}_t$. Consider a data matrix with the most recent n observation available at time t, which we denote by

$$\mathbf{X}_t^{(n)} = \begin{bmatrix} \boldsymbol{x}_{t-n+1} & \boldsymbol{x}_{t-n+2} & \cdots & \boldsymbol{x}_t \end{bmatrix}. \tag{2.20}$$

Following the analysis in Section 2.1, at time t the sample covariance matrix

$$\mathbf{S}_t^{(n)} = \frac{1}{n-1} \mathbf{X}_t \left(\mathbf{1}_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n' \right) \mathbf{X}_t' \tag{2.21}$$

has an eigenvalue decomposition as in equation (2.15). This yields an empirical eigenspectrum $\boldsymbol{l}_t = [l_{1t}, \cdots, l_{dt}]'$ and replacing l_i with l_{it} in (2.17)-(2.18) we get $\hat{\Omega}_t$ and $\hat{\omega}_{it}$, respectively. Finally,

$$\hat{\phi}_{kt} = \frac{\hat{\omega}_{kt}}{\hat{\Omega}_t}. \tag{2.22}$$

2.1.3 Eigenspectrum Analysis versus Average Pearson Correlation

For the purpose of assessing the integration between two financial assets, one may naturally question the difference between the framework above and a simple averaging of the elements of the correlation matrix. Obviously strong correlations of opposite signs can result in a low average correlation coefficient. However, even computing the average absolute value of correlation coefficients can be disadvantageous. The PCA approach provides a more granular view of the

structure of asset co-movements. Specifically, we can see the concentration of variance captured by eigenvalues of different rank and this can be meaningful information, especially when different eigenvectors have different economic interpretations. For example, the dominant eigenvector has been identified as influence of the entire market and subsequent eigenvectors can explain clusters of stocks with similar return dynamics, which have been interpreted as industry effects, for example [1, 45, 60, 66]. In addition, we may be interested in a measure of co-movement that takes into account the different return volatilities of assets. In this case, the relative magnitude of eigenvalues of the covariance matrix will be a better measure than the average correlation coefficient [43].

Generally, there are various ways in which one can define and quantify the strength of co-movement in assets. In addition to the two approaches discussed above, factor models have been used to measure the strength of economic integration [69]. For example, consider a two-factor model

$$r_1(t) = \alpha_1 + \beta_{11} f_1(t) + \beta_{12} f_2(t) + \epsilon_1(t) \tag{2.23}$$

$$r_2(t) = \alpha_2 + \beta_{21} f_1(t) + \beta_{22} f_2(t) + \epsilon_2(t), \tag{2.24}$$

where at time t, $r_i(t)$ is the return on stock i, $f_j(t)$ is the j-th factor and β_{ij} is the sensitivity coefficient of returns on stock i to factor j, with $i = 1, 2, j = 1, 2$ and $t = 1, \ldots, T$. There is also an idiosyncratic component $\epsilon_i(t)$ for each stock. If we assume that $\epsilon_1(t) = \epsilon_2(t) = 0, \forall t$ then the two stocks can be considered perfectly integrated since their return dynamics are fully defined by common two factors. Assuming, without loss of generality, that $\mu_i = \frac{1}{T}\sum_{t=1}^{T} r_i(t) = 0$, we have

$$\text{Corr}\left[r_1(t), r_2(t)\right] = \frac{\text{Cov}\left[r_1(t), r_2(t)\right]}{\sigma_1 \sigma_2} \tag{2.25}$$

$$= \frac{\sum_{t=1}^{T} r_1(t) r_2(t)}{\sqrt{\sum_{t=1}^{T} r_1^2(t)} \sqrt{\sum_{t=1}^{T} r_2^2(t)}} \tag{2.26}$$

$$\leq 1 \tag{2.27}$$

and from the well-known Cauchy-Schwarz inequality, equation (2.27) holds with equality if and only if $r_1(t)$ and $r_2(t)$ are linearly dependent. That is, the Pearson correlation coefficient represents perfect integration if and only if $\beta_{1j} = k\beta_{2j}$, so that the factor sensitivities of the two stocks are proportional. If this condition is not satisfied, then Pearson correlation will be less than one, representing imperfect integration. Perfect integration in the PCA approach would reflect in a single non-zero eigenvalue and this also happens if and only if $r_1(t)$ and $r_2(t)$ are linearly dependent. Concluding that simple correlation can be a poor measure of economic integration, the authors in [69] propose to use the R-squared statistic from regressions of returns on common factors in order to assess the strength of integration. They show that this technique points to increased integration between international stock markets due to globalization, while correlations

decrease over the same period. It should be noted, however, that any analysis of integration of asset returns in a linear framework likely fails to capture their true dynamics. Non-linear features of empirical returns are well-documented and there is vast literature on modeling techniques that address this issue (See [32, Section 1.2], for example).

2.2 The Distribution of Eigenvalues

We will now discuss the distributional properties of correlation matrix spectra. Some work on the joint distribution of the eigenvalues has been done in multivariate statistical analysis. However, most results on the limiting behaviour of eigenvalues come from Random Matrix Theory. In this section we present theoretical results on the distribution of eigenvalues.

2.2.1 Random Matrices

A random matrix is a matrix whose elements are random variables. We shall denote random matrices by calligraphic letters to distinguish them from those with deterministic elements. Thus, a random matrix, $\mathcal{X}$, has $X_{ij} \sim F$ as its (i,j)-th element, for some distribution F. The corresponding realizations are $\mathbf{X}$ and x_{ij}. Random matrices are studied as ensembles, each of which includes matrices that usually form an algebraic group [11]. We list some of the most studied ensembles and their construction [31] below.

- **Gaussian orthogonal ensemble**: symmetric matrices that can be written as $\frac{(\mathcal{X}+\mathcal{X}')}{2}$ where $\mathcal{X}$ is $n \times n$ with i.i.d. standard real Normal elements.
- **Gaussian unitary ensemble**: Hermitian matrices that can be written as $\frac{(\mathcal{X}+\mathcal{X}^H)}{2}$ where $\mathcal{X}$ is $n \times n$ with i.i.d. standard complex Normal elements. We use $\mathcal{X}^H$ to denote the Hermitian transpose of $\mathcal{X}$.
- **Gaussian symplectic ensemble**: self-dual matrices that can be written as $\frac{(\mathcal{X}+\mathcal{X}^D)}{2}$ where $\mathcal{X}$ is $n \times n$ with i.i.d. standard quaternion Normal elements. We use $\mathcal{X}^D$ to denote the dual transpose of $\mathcal{X}$.
- **Wishart ensemble**: symmetric matrices that can be written as $\mathcal{X}\mathcal{X}^T$, where $\mathcal{X}$ is $d \times n$ with i.i.d. standard Normal elements that can be real, complex or quaternion. We let $\mathcal{X}^T$ denote $\mathcal{X}'$, $\mathcal{X}^H$ or $\mathcal{X}^D$ appropriately.

Of considerable importance to multivariate statistical analysis is the study of Wishart matrices. Next, we will review the Wishart distribution and results relating to the eigenspectrum of Wishart matrices.

2.2.2 Wishart distribution

Wishart matrices arise most prominently in the analysis of sample covariance matrices. Named after John Wishart, who first computed its joint element density, the Wishart distribution is a generalization of the chi-squared distribution to higher dimension.

Definition 10 (Wishart Distribution)**.** Let $\mathcal{M} \in \mathbb{R}^{d\times d}$ be a symmetric random matrix such that

$$\mathcal{M} = \mathcal{X}\mathcal{X}' \tag{2.28}$$

where

$$\mathcal{X} = \begin{bmatrix} \boldsymbol{X}_1 & \cdots & \boldsymbol{X}_n \end{bmatrix} \tag{2.29}$$

and $\boldsymbol{X}_i \sim N_d(\mathbf{0}, \boldsymbol{\Sigma})$ independently for $i = 1, \ldots, n$. Then $\mathcal{M}$ has a *Wishart distribution with covariance* $\boldsymbol{\Sigma}$ *and* n *degrees of freedom*, which we denote by

$$\mathcal{M} \sim W_d(n, \boldsymbol{\Sigma}). \tag{2.30}$$

Recall that $\mathcal{X}'\mathcal{X}$ is positive-semidefinite and thus $\boldsymbol{c}'\mathcal{M}\boldsymbol{c} \geq \mathbf{0}, \forall \boldsymbol{c} \in \mathbb{R}^d$. Now, a Wishart matrix is nonsingular if and only if $n \geq d$ [57, p. 82]. Under this condition, $\mathcal{M}\boldsymbol{c} \neq \mathbf{0}$ so that $\boldsymbol{c}'\mathcal{M}\boldsymbol{c} > \mathbf{0}$ for every nonzero $\boldsymbol{c}$. In other words, a Wishart matrix is positive-definite if and only if $n \geq d$. Only when this condition holds does the Wishart matrix have a density function, as formalized in the following theorem from Muirhead [57].

Theorem 1. *Assume that* $\mathcal{M} \sim W_d(n, \boldsymbol{\Sigma})$ *with* $n \geq d$*. Then* $\mathcal{M}$ *has the density function*

$$f_{\mathcal{M}}(\mathbf{M}; n, d, \boldsymbol{\Sigma}) = \frac{1}{2^{\frac{dn}{2}}\Gamma_d(\frac{n}{2})\left(\det \boldsymbol{\Sigma}\right)^{\frac{n}{2}}} \operatorname{etr}\left(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\mathbf{M}\right)\left(\det \mathbf{M}\right)^{\frac{n-d-1}{2}}, \tag{2.31}$$

where $\Gamma_d(\cdot)$ *denotes the multivariate gamma function.*

Having introduced the Wishart distribution, we can now view the sample covariance matrix as a random matrix $\mathcal{S}$. Recall that, unlike $\mathbf{S}$, as defined in (1.9), the elements of $\mathcal{S}$ are treated as random variables. Then, we have that $\mathcal{S} \sim W_d\left(n-1, \frac{\boldsymbol{\Sigma}}{n-1}\right)$. Thus, the joint distribution of its elements exists if and only if $n \geq d$, in which case it is given by

$$f_{\mathcal{S}}(\mathbf{S}; n, d, \boldsymbol{\Sigma}) = \frac{\left(\frac{n}{2}\right)^{\frac{dn}{2}}}{\Gamma_d(\frac{n}{2})\left(\det \boldsymbol{\Sigma}\right)^{\frac{n}{2}}} \operatorname{etr}\left(-\frac{1}{2}n\boldsymbol{\Sigma}^{-1}\mathbf{S}\right)\left(\det \mathbf{S}\right)^{\frac{n-d-1}{2}}. \tag{2.32}$$

The eigenspectrum of $\mathcal{S}$ is of interest both in the general study of random matrices, not least due to its importance in multivariate statistics, and, as per Section 2.1, in our study of co-movement in multivariate series. Modeling the sample covariance matrix as a random matrix

allows for analysis that is based on tools from Random Matrix Theory. However, we shall first present what is known about the distribution of eigenvalues from multivariate analysis. Begin with the following theorem from James [38].

Theorem 2 (Joint Density of Eigenvalues)**.** *Let $\mathcal{S}$ be a random matrix distributed as $W_d(N, \frac{\mathbf{\Sigma}}{N})$, with $N = n - 1 \geq d$. Define $\mathcal{L} = \operatorname{diag}(l_1, \ldots, l_d)$ to be a diagonal matrix of the eigenvalues of $\mathcal{S}$ and $\mathcal{U} \in \mathcal{O}(d)$ such that $\mathcal{S} = \mathcal{U}\mathcal{L}\mathcal{U}'$. Then the joint density of the eigenvalues $l_1 > \ldots > l_d > 0$ is*

$$f_{\mathcal{L}}(l_1 = x_1, \ldots, l_d = x_d; N, d, \mathbf{\Sigma}) = \left(\frac{N}{2}\right)^{\frac{dN}{2}} \frac{\pi^{\frac{d}{2}^2} (\det \mathbf{\Sigma})^{-\frac{N}{2}}}{\Gamma_d\left(\frac{N}{2}\right)\Gamma_d\left(\frac{d}{2}\right)} \prod_{i=1}^{d} x_i^{\frac{N-d-1}{2}} \prod_{i<j}^{d} (x_i - x_j) \cdot \int_{\mathcal{O}(d)} \operatorname{etr}\left(-\frac{N}{2}\mathbf{\Sigma}^{-1}\mathbf{U}\mathbf{X}\mathbf{U}'\right) d\mathbf{U}, \tag{2.33}$$

where $\mathbf{X} = \operatorname{diag}(x_1, \ldots, x_d)$ and $(d\mathbf{U})$ is the normalized invariant measure on $\mathcal{O}(d)$.

The integral term in (2.33) is in general difficult to evaluate analytically. In the null case, $\mathbf{\Sigma} = \lambda \mathbf{I}_d$, using the facts that

$$\int_{\mathcal{O}(d)} d\mathbf{U} = 1 \tag{2.34}$$

and

$$\operatorname{tr}(\mathbf{U}\mathbf{X}\mathbf{U}') = \operatorname{tr}(\mathbf{X}), \tag{2.35}$$

we have

$$\int_{\mathcal{O}(d)} \operatorname{etr}\left(-\frac{N}{2}\mathbf{\Sigma}^{-1}\mathbf{U}\mathbf{X}\mathbf{U}'\right) d\mathbf{U} = \int_{\mathcal{O}(d)} \operatorname{etr}\left(-\frac{N}{2\lambda}\mathbf{U}\mathbf{X}\mathbf{U}'\right) d\mathbf{U} \tag{2.36}$$

$$= \operatorname{etr}\left(-\frac{N}{2\lambda}\mathbf{X}\right) \int_{\mathcal{O}(d)} d\mathbf{U} \tag{2.37}$$

$$= \exp\left(-\frac{N}{2\lambda}\sum_{i=1}^{d} x_i\right). \tag{2.38}$$

Thus the joint density of the eigenvalues becomes

$$f_{\mathcal{L}}(l_1 = x_1, \ldots, l_d = x_d; N, d, \lambda\mathbf{I}_d) = \left(\frac{N}{2\lambda}\right)^{\frac{dN}{2}} \frac{\pi^{\frac{d}{2}^2}}{\Gamma_d\left(\frac{N}{2}\right)\Gamma_d\left(\frac{d}{2}\right)} \prod_{i=1}^{d} x_i^{\frac{N-d-1}{2}} \prod_{i<j}^{d} (x_i - x_j) \cdot \exp\left(-\frac{N}{2\lambda}\sum_{i=1}^{d} x_i\right). \tag{2.39}$$

However, the non-null case is much more involved. James [38] obtains an infinite series represen-

tation for a general $\mathbf{\Sigma}$ in terms of a hypergeometric function of matrix arguments:

$$\int_{\mathcal{O}(d)} \text{etr}\left(-\frac{1}{2}N\mathbf{\Sigma}^{-1}\mathbf{U}\mathbf{X}\mathbf{U}'\right)d\mathbf{U} = {}_0F_0{}^{(d)}(-\frac{N\mathbf{X}}{2}, \mathbf{\Sigma}^{-1}). \tag{2.40}$$

$$\tag{2.41}$$

See [57, pp. 258] for a thorough introduction.

Using similar tools an exact expression of distribution of the largest eigenvalue, l_1, is (see [57, 21])

$$P(l_1 \leq x) = \frac{\Gamma_d\left(\frac{d+1}{2}\right)}{\Gamma_d\left(\frac{d+N}{2}\right)}\det\left(\frac{N}{2}x\mathbf{\Sigma}^{-1}\right)^{\frac{N}{2}} {}_1F_1{}^{(d)}\left(\frac{N}{2}; \frac{N+d}{2}; -\frac{N}{2}x\mathbf{\Sigma}^{-1}\right). \tag{2.42}$$

The distribution of the largest eigenvalue is useful in various applications. For example, in the estimation of a sparse mean vector, the maximum of d i.i.d Gaussian noise variables is of interest [41]. In addition, PCA is often used as a dimensionality reduction tool in which extreme eigenvalues play a role in the choice of how many PCs to retain. Most importantly for this work, the relative energy of the largest eigenvalue produced from asset returns can shed light on the interconnectedness of a market, and hence its propensity to experience systemic shocks.

There are issues in practical applications of the density representations mentioned so far. Firstly, the true covariance matrix, $\mathbf{\Sigma}$, is a required parameter that is usually not known. Secondly, the non-null case of $\mathbf{\Sigma}$ results in expressions that are difficult to evaluate analytically. Hypergeometric functions involve infinite series that converge very slowly for large d, and so even numerical solutions are typically inefficient. We proceed with reviewing some results from random matrix theory that are more useful in this sense.

2.2.3 Results from Random Matrix Theory

Random Matrix Theory (RMT) addresses the properties of large matrices whose entries are random variables. The study of random matrices has focused on their eigenvalues as early as the 1920s with the work of Wishart [87]. A couple of decades later the field has become prominent in nuclear physics where dynamic systems were approximated by discretization, leading to matrices of very large dimensions. This motivated interest in the limiting behaviour of the eigenvalues and indeed the pioneering work of Wigner [86] and Marčenko and Pastur [53] is concerned with applications in physics. Significant improvements in computing have since proliferated studies of high dimensional data and, nowadays, large matrices are often seen in many fields. Results from RMT have been used in such areas as:

1. Wireless communication: d input channels and n output channels. See [83], for example.
2. Climate studies: d measurement locations and n time points. See [68], for example.

3. Financial data: d is the number of securities and n is the number of return observations. Examples are [35, 45].

Much of RMT is concerned with the spectral distribution of infinitely large matrices. Specifically, as the size of the matrix grows infinitely large, the eigenvalue behaviour is described in terms of the limiting empirical spectral distribution.

Definition 11 (Empirical Spectral Distribution)**.** Let $\mathbb{1}_{[A]}$ be the indicator function for event A. The *empirical spectral distribution* (*ESD*) of a square matrix $\mathbf{X}$ is defined by

$$F^{\mathbf{X}}(x) = \frac{1}{d}\sum_{i=1}^{d} \mathbb{1}_{[l_i \leq x]}. \tag{2.43}$$

Put otherwise, $F^{\mathbf{X}}(x)$ equals the proportion of eigenvalues not greater than x.

If the ESD of a matrix converges to a non-random distribution F as its dimension increases to infinity, then F is said to be the *limiting spectral distribution* (LSD). Analysis of the LSD of large dimensional random matrices is a central theme in RMT and is the subject of the most famous results, such as Wigner's *semi-circle law* [86] and its analog for other random matrix ensembles. Bai and Silverstein [8] provide a recent comprehensive reference.

Suppose, for instance, that $\mathcal{X}$ is such that its elements are independently distributed random variables $X_{ij} \sim N(0,1)$. The population covariance is given by $\mathbf{I}_d$ and we have that $\lambda_1 = \lambda_2 = \ldots = \lambda_d = 1$. However, a Monte Carlo simulation of randomly drawn $\mathcal{X}$ with $n = d = 10$ produces[2] the ordered eigenvalues $l_1 > l_2 > \ldots > l_{10}$ with systematic errors, as summarized in Table 2.1. The implication of this argument is that we cannot expect the empirical eigenvalues to have equal energy even if the population eigenvalues are in fact equal. This phenomenon is sometimes described as greater spacing between sample eigenvalues than their population counterparts.

Statistic	l_1	l_2	l_3	l_4	l_5	l_6	l_7	l_8	l_9	l_{10}
Observed	3.05	2.25	1.67	1.21	0.84	0.53	0.29	0.13	0.03	0.00
Error	2.05	1.25	0.67	0.21	-0.16	-0.47	-0.71	-0.87	-0.97	-1.00

Table 2.1: Ordered eigenvalues of realizations of $\mathbf{S}$ with $n = d = 10$, averaged over 1000 simulations.

The next result demonstrates that the distribution of the eigenvalues as $n, d \to \infty$ converges to a non-random distribution. For the case when $\mathcal{X}$ has real elements, the LSD is known as the Marčenko-Pastur (MP) law [53].

[2] After subtracting the mean and standardizing.

Theorem 3 (Marčenko-Pastur law)**.** *Let $\mathcal{X} \in \mathbb{R}^{d\times n}$ be a matrix whose columns are d-dimensional random variables $\boldsymbol{X}_j = [X_{1j} \ \dots \ X_{dj}]'$ satisfying:*

1. *Independence: the elements X_{ij} are i.i.d. $\forall i, j$.*
2. *Zero mean: $\mathbb{E}(X_{ij}) = 0 \quad \forall i, j$.*
3. *Constant and uniform variance: $\mathbb{E}(X_{ij}^2) = \sigma^2 \quad \forall i, j$.*
4. *Finite fourth moment: $\mathbb{E}(|X_{ij}|^4) < \infty \quad \forall i, j$.*
5. *Asymptotic aspect ratio: $n, d \to \infty$ such that $\frac{d}{n} \to y \leq 1$.*

Then, the ESD of the sample covariance matrix, $\mathcal{S}$, converges in probability to the Marčenko-Pastur distribution with probability density function given by

$$f_y(x) = \frac{\sqrt{(b-x)(a-x)}}{2\pi\sigma^2 yx}\mathbb{1}_{[a,b]}(x), \tag{2.44}$$

where $a = \sigma^2(1-\sqrt{y})^2$ and $b = \sigma^2(1+\sqrt{y})^2$.

The assumption of independence in the Marčenko-Pastur law regards both the *row-wise* and *column-wise* structure of $\mathcal{X}$. In the case of a d-dimensional time series, $\{\boldsymbol{X}_t\}$, this is called *contemporaneous* and *temporal* structure, respectively. Multivariate data without any structure, such as white noise, will have a LSD with the probability density function given in equation (2.44). However, in many practical setting the dependence structure between the entries of $\boldsymbol{X}_t$ is not as trivial and the LSD in such cases will be different. Some extensions of the Marčenko-Pastur Law have addressed contemporaneous dependence, i.e. dependence between the components $X_i, i = 1, \dots, d$. Silverstein [75] considers matrices of the form $\mathcal{Y} = \mathcal{T}_p^{\frac{1}{2}}\mathcal{X}$ where $\{\mathcal{T}_p\}$ is a sequence of Hermitian non-negative definite matrices. A strong LSD is established and a characteristic equation for its Stieltjes transform is given for $\frac{1}{n}\mathcal{Y}\mathcal{Y}'$ under the assumption that $\{\mathcal{T}_p\}$ is bounded in spectral norm and its spectral density converges in distribution to a non-random density. More recently, Bai and Zhou [7] extended the result to a more general $\{\mathcal{T}_p\}$ that satisfies a mild moment condition. There have also been studies that consider temporal dependence using linear models. Jin et al. [39] establish the existence of a LSD for large sample covariance matrices generated by a vector autoregressive moving average (VARMA) process. An explicit form for the probability density function is provided for VAR(1) and VMA(1) models. Yao [89] studied a similar problem and determined the LSD and its Stieltjes transform assuming a general linear process. Most recently, Liu et al. [49] extend the Marčenko-Pastur law to multivariate MA(∞) models and find the LSD of the symmetrized lagged autocovariance matrix in terms of its Stieltjes transform. Davis et al. [24] also consider sample covariance matrices resulting when the rows of $\boldsymbol{X}_t$ are copies of some time-dependent linear process. Their main result states that the joint distribution of the first k eigenvalues converges to a Poisson point process.

The distribution of the largest eigenvalue in the Wishart ensemble is of particular importance in applications such as hypothesis testing and signal detection. While an exact expression for this had been known, and is given in (2.42), its hypergeometric function factor renders it intractable in many practical settings. Surveys by Pillai [64, 65] contain additional discussions on challenges in obtaining marginal distributions from the joint distribution of eigenvalues of Wishart matrices. In some cases, RMT techniques have been useful in this context. Johnstone [41] establishes that the limiting distribution of the largest eigenvalue obeys the *Tracy-Widom law* in the following theorem.

Theorem 4 (Johnstone's Theorem)**.** *Let $\mathcal{X}$ be a $d \times n$ random matrix with i.i.d. elements $X_{ij} \sim N(0,1)$. Let l_1 be the largest sample eigenvalue of $\mathcal{M} = \mathcal{X}\mathcal{X}'$ and define the centre and scaling constants*

$$\mu_{dn} = \left(\sqrt{n-1} + \sqrt{d}\right)^2 \tag{2.45}$$

$$\sigma_{dn} = \left(\sqrt{n-1} + \sqrt{d}\right)\left(\frac{1}{\sqrt{n-1}} + \frac{1}{\sqrt{d}}\right)^{\frac{1}{3}}. \tag{2.46}$$

If $d, n \to \infty$ such that $\dfrac{d}{n} \to y \leq 1$, then

$$\frac{l_1 - \mu_{dn}}{\sigma_{dn}} \xrightarrow{\mathfrak{D}} W_1 \sim F_{TW}^{(1)} \tag{2.47}$$

where the convergence is to the Tracy-Widom law of order 1, which has the distribution function

$$F_{TW}^{(1)}(y) = \exp\left(-\frac{1}{2}\int_y^\infty q(x) + (x-y)q^2(x)dx\right) \qquad y \in \mathbb{R} \tag{2.48}$$

where $q(x)$ solves the Painlevé II differential equation

$$\begin{aligned} q''(x) &= xq(x) + 2q^3(x), \\ q(x) &\sim \mathrm{Ai}(x) \text{ as } x \to \infty, \end{aligned} \tag{2.49}$$

and $\mathrm{Ai}(x)$ denotes the Airy function.

The distribution in (2.48) was obtained by Tracy and Widom [81] as the limiting law of the largest eigenvalue of an $n \times n$ Gaussian symmetric matrix. However, Johnstone's result states that it also applies to Wishart matrices if l_1 is standardized by the constants in (2.45) and (2.46). Evaluating the Painlevé II differential equation, and hence the Tracy-Widom distribution, requires numerical approximation techniques such as in Edelman and Persson [30]. Bejan [11] provides tables with evaluations and p-values that are analogous to the traditional tables for the Student's t-distribution, for example.

2.3 Stylized Properties of Eigenvalue-based Time Series

Our objective here is to expand on the notion of a spectrum-based time series as presented in Section 2.1.2. Ultimately, we are interested in performing a sequential analysis of the strength of contemporaneous dependence in multivariate time series. We recap the setting: consider, at time $t \geq n$, a data matrix $\mathbf{X}_t^{(n)}$ of the most recent n observations $\boldsymbol{x}_{t-n+1}, \ldots, \boldsymbol{x}_t$ from a d-variate time series, $\{\boldsymbol{X}_t\}$. The corresponding sample covariance matrix is $\mathbf{S}_t^{(n)}$ and its eigenvalues, $\boldsymbol{l}_t$, can be used to construct a single observation of the statistics $\hat{\omega}_{it}$ or $\hat{\phi}_{kt}$ as defined in equations (2.18)-(2.19). We can also generate series $\{\boldsymbol{l}_t\}$, $\{\hat{\omega}_{it}\}$ and $\{\hat{\phi}_{kt}\}$ by applying PCA sequentially on a fixed-size rolling window of size n.

For clarification, let us outline this procedure for an i.i.d series of $\boldsymbol{X}_t \sim N_4(\mathbf{0}, \mathbf{I}_4)$. We generate n arbitrarily ordered (index-wise) observations which we assume is information available at the start of sampling. We compute an initial $\mathbf{S}_0^{(n)}$ and its eigenvalues, $\boldsymbol{l}_0$. Upon arrival of the next observation, we truncate the oldest observation, leaving the most recent n points, with which $\mathbf{S}_1^{(n)}$ and $\boldsymbol{l}_1$ are computed. Let this process continue and the series of interest will be generated. Figure 2.1 shows the relative magnitude series of eigenvalues produced from a Monte Carlo simulation of 1000 sequential spectra using $n = 100$. Each eigenvalue accounts for approximately a quarter of the variance. In Figure 2.2 we plot cumulative explained variance. Recall that $\hat{\phi}_k$ is simply the sum of the first k eigenvalues, normalized by the sum of all eigenvalues. If we choose to standardize the data to have unit variance then the eigenvalues would simply be scaled down by d, the number of variates.

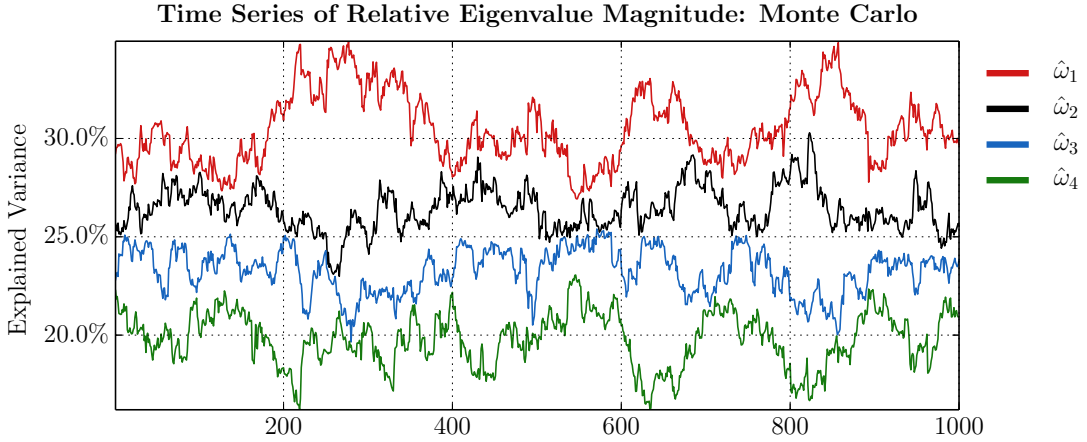

Figure 2.1: Normalized eigenvalues of a sample covariance matrix $\mathbf{S}_t^{(n)}$ of $\boldsymbol{X}_t \sim N_4(\mathbf{0}, \mathbf{I_4})$, with an initial sample of $n = 100$ rolled over 1000 sequential observations.

It is useful to compare the empirical eigenvalues to the theoretical ones. In the above experiment we obtain a mean value of 1.225 for l_1, whereas the theoretical upper support of the Marčenko-Pastur for an aspect ratio $y = \dfrac{4}{100}$ is 1.44. If we increase both d and n by a factor of 100, l_1 comes at a mean value of 1.39. In fact, the limits of the smallest and largest eigenvalues

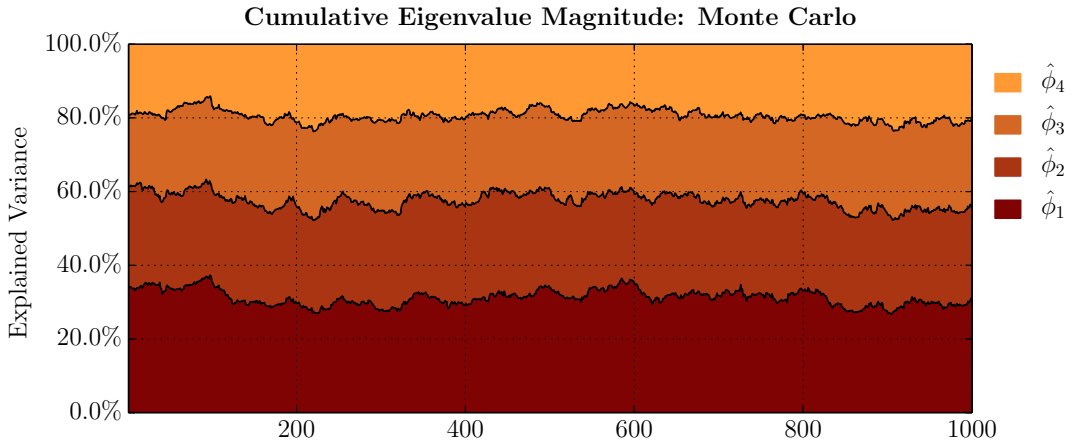

Figure 2.2: Cumulative normalized eigenvalues of a sample covariance matrix $\mathbf{S}_t^{(n)}$ of $\boldsymbol{X}_t \sim N_4(\mathbf{0}, \mathbf{I_4})$, with an initial sample of $n = 100$ rolled over 1000 sequential observations.

have been shown (see Geman [34] and Silverstein [74]) to converge to the lower and upper edges of the support of the Marčenko-Pastur density, denoted by a and b in Theorem 3, respectively. Formally, we have

$$l_1 \to (1 + \sqrt{y})^2, \; a.s. \tag{2.50}$$

$$l_d \to (1 - \sqrt{y})^2, \; a.s. \tag{2.51}$$

where the almost sure convergence is to be interpreted as occurring with probability one as $d, n \to \infty$. Thus, comparing empirical extreme eigenvalues with their theoretical counterparts can give information about the structure in the data. If the data is white noise, as is the case underlying the Marčenko-Pastur law, then we would expect empirics to be close to theory. In contrast, a relatively strong structure in the data would result in a significantly larger dominant eigenvalue than predicted by RMT.

2.3.1 Autocorrelation

Clearly, each $\{\hat{\omega}_{it}\}, i = 1, \ldots, d$, above is not serially independent because their values are obtained using a rolling window of observations. In our construction, the last $n - 1$ columns of $\mathbf{X}_t^{(n)}$ are the same as the first $n - 1$ columns of $\mathbf{X}_{t+1}^{(n)}$. The window size, n, acts as a smoothing parameter in the sense that larger n result in higher autocorrelations and smaller variance in the eigenvalue series. However, the window size also affects the mean values through the dependence of the eigenvalue on the aspect ratio $y = \dfrac{d}{n}$. These dynamics are illustrated in Figure 2.3, where we plot $\{\hat{\omega}_{1t}\}$ and its autocorrelation functions for simulations using windows of varying length.

It can be seen from the figure that as n increases, the autocorrelation decays slower. Of course, serial dependence could be statistically eliminated if we lagged our sampling, such that there is no

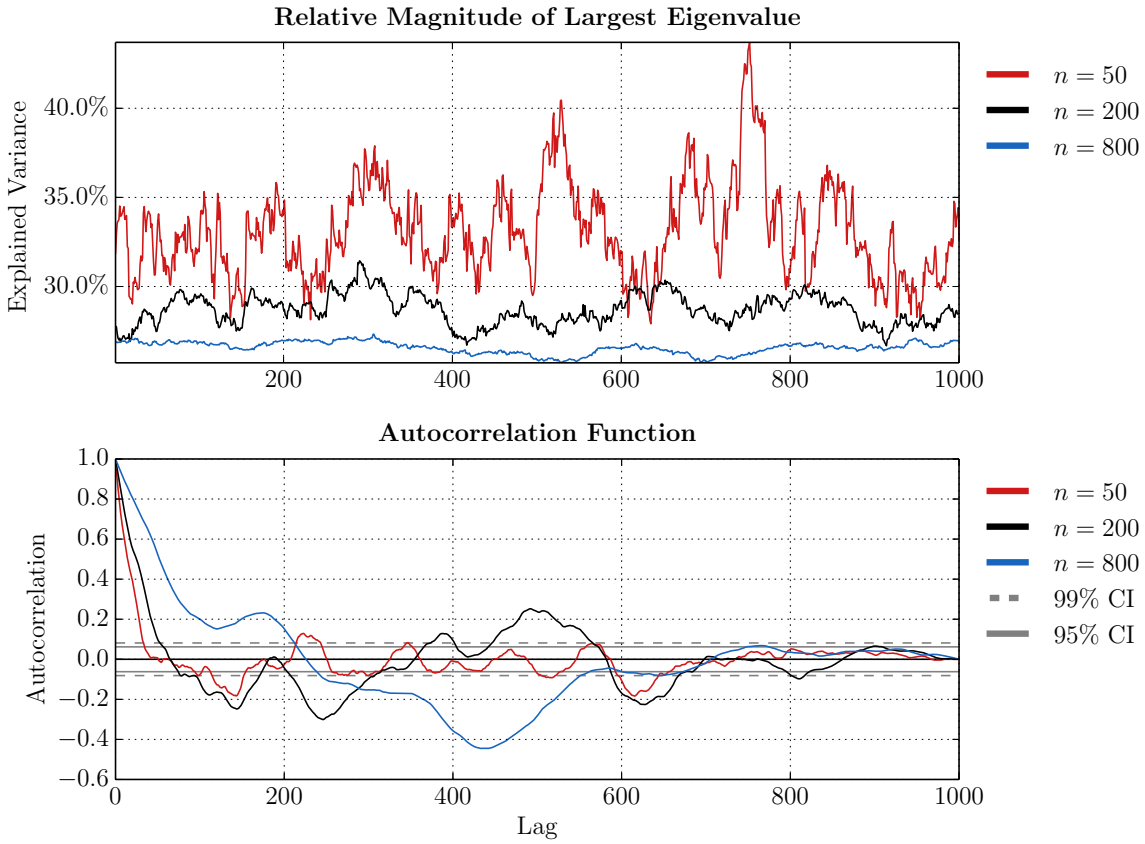

Figure 2.3: Top – cumulative normalized eigenvalues of a sample covariance matrix $\mathbf{S}_t^{(n)}$ of $\boldsymbol{X}_t \sim N_4(\mathbf{0}, \mathbf{I_4})$, with samples of varying window length. Bottom – autocorrelation functions and confidence intervals for each series in the top panel.

overlap between successive data matrices $\boldsymbol{X}_t$. However, a more common approach is to difference the series such that each observation represents increments $\Delta\hat{\omega}_{1t} = \hat{\omega}_{1t} - \hat{\omega}_{1t-1}$. In Figure 2.4 we plot the first difference of the resulting $\{\hat{\omega}_{1t}\}$ when $n = 50$. Note that the autocorrelation values now fall within the confidence intervals around zero correlation. This is rather intuitive as this differenced series reflects changes in the amount of variance explained by the largest eigenvalues estimated from successive i.i.d. observations of the data.

2.3.2 Stationarity

Results from Section 2.2 on the distribution of eigenvalues provide justification for the claim of eigenvalue series stationarity under certain conditions. Consider first the relative magnitude of single eigenvalues $\hat{\omega}_{it}, i = 1, \ldots, d$. The stationarity of each depends on that of l_{it}. We know that in the case of a d-variate i.i.d. Gaussian process, the distribution of the largest eigenvalue exists and is given by equation (2.42). Asymptotic distributions are given by Anderson [2] for fixed d

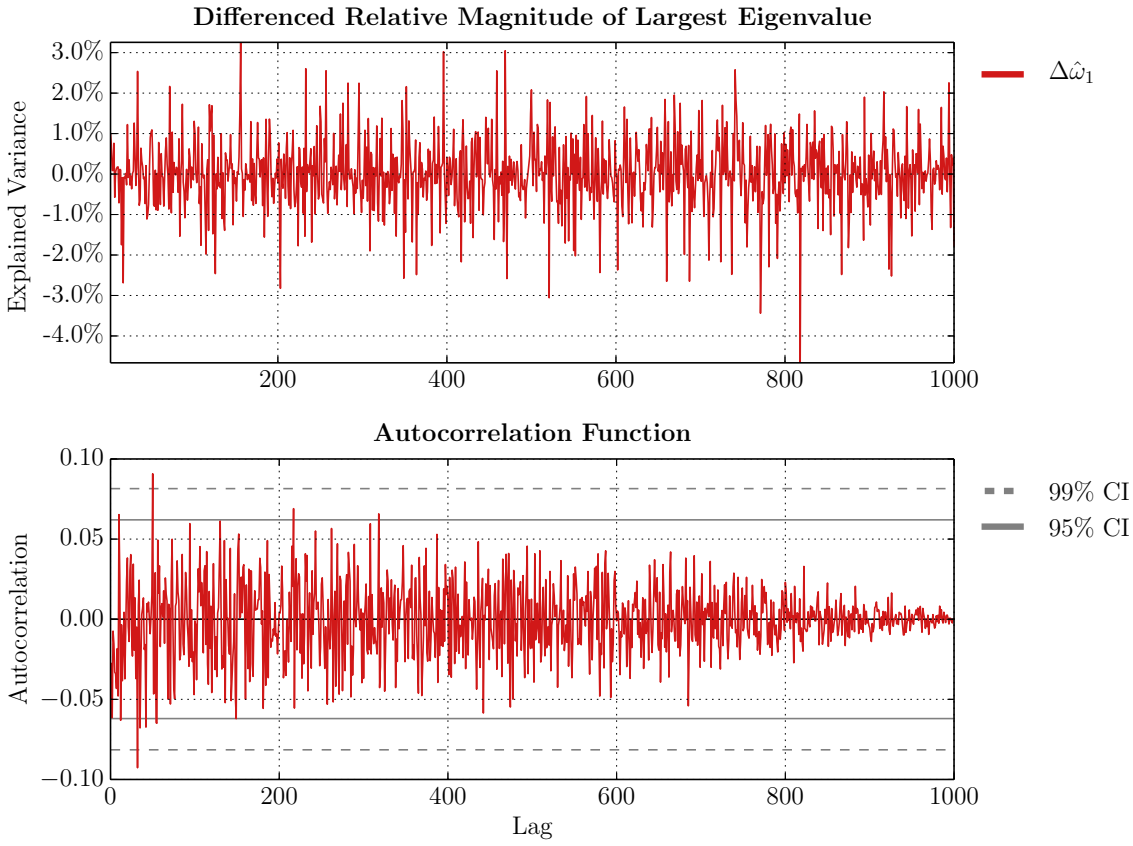

Figure 2.4: Top – first difference of $\hat{\omega}_{1t}$ generated with $n = 50$. Bottom – Autocorrelation function with confidence intervals for the series in the top panel.

and $n \to \infty$ in the form

$$\sqrt{\frac{n}{2}}\left(\frac{l_1}{n} - 1\right) \xrightarrow{\mathfrak{D}} N(0,1), \tag{2.52}$$

and by Johnstone [41] as in Theorem 4. Denote this by $F_{l_1}(l_1 \leq x; n, d, \mathbf{\Sigma})$ to emphasize the parameter set of the distribution. Clearly the conditions under which $\{l_{1t}\}$ is stationary depend on the underlying time series and the sampling procedure which produces each l_{1t}. Specifically, if the underlying process is at least weakly-stationary, then taking disjoint samples $\mathbf{X}_t^{(n)}$ will result in i.i.d realizations l_1. To the best of our knowledge, there are no results on non-extreme eigenvalues.

However, in empirical applications, the rolling-window fashion by which the eigenvalues are calculated results in some dependency. Thus, we conduct three tests on $\hat{\omega}_{it}$ and $\Delta\hat{\omega}_{it}$ to investigate their stationarity: (1) Augmented Dickey-Fuller (ADF) test [26], (2) Phillips-Perron (PP) test [63] and (3) Kwiatkowski, Phillips, Schmidt and Shin (KPSS) test [44]. The null hypotheses for

Test	$\hat{\omega}_i$				$\Delta\hat{\omega}_i$			
	$i=1$	$i=2$	$i=3$	$i=4$	$i=1$	$i=2$	$i=3$	$i=4$
Reject H_0^{ADF}	0%	0%	0%	0%	100%	100%	100%	100%
Reject H_0^{PP}	0%	0%	0%	0%	100%	100%	100%	100%
Do not reject H_0^{KPSS}	0%	0%	0%	0%	98%	100%	92%	94%

Table 2.2: stationarity tests for 1000 simulated series with $n = 100$ and $d = 4$

Test	$\hat{\phi}_k$			$\Delta\hat{\phi}_k$		
	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$
Reject H_0^{ADF}	0%	0%	0%	100%	100%	100%
Reject H_0^{PP}	0%	0%	0%	100%	100%	100%
Do not reject H_0^{KPSS}	0%	0%	0%	92%	100%	96%

Table 2.3: stationarity tests for 1000 simulated series with $n = 100$ and $d = 4$.

these tests are

$$H_0^{ADF} = H_0^{PP} : \text{the series has a unit root,} \tag{2.53}$$

$$H_0^{KPSS} : \text{the series does not have a unit root.} \tag{2.54}$$

A unit root is a feature of nonstationary series, so that a rejection of the hypothesis in (2.53) for a given series is to be interpreted that it is stationary. On the contrary, the opposite conclusion is drawn from the KPSS test if there is sufficient evidence to reject its null hypothesis, H_0^{KPSS}. Results for our tests are summarized in Table 2.2, where the relative magnitude and changes in relative magnitude for each eigenvalue are taken into consideration. The tests consistently conclude that $\hat{\omega}_{it}$ is not stationary since, in 1000 simulations, there are no rejections of the unit root hypothesis. However, the differenced series is found to be stationary by all tests in the vast majority of simulations. Table 3 summarizes results from similar tests for the cumulative sums $\hat{\phi}_{kt} = \sum_{i=1}^{k} \hat{\omega}_{it}, k = 1, \ldots, d-1$. As before, the series are found to be stationary only after differencing.

Chapter 3

Empirical Study

In this chapter we study the connection between systemic risk and asset return correlations using the spectral analysis scheme presented in the previous chapter. There are several works in extant literature that relate to ours in the use of spectral analysis and RMT to study the level of correlation between financial assets and its implications on risk management. In particular, a recursive PCA scheme has been implemented to gain insight into the temporal evolution of the interconnectedness of financial assets and its relationship with systemic shocks (see [14, 43, 54], for examples). However, so far such an analysis had only been made to illustrate a relationship visually. Here, we expand on this approach by studying the properties of correlation matrix eigenvalues in empirical data, and what can be statistically inferred from the eigenvalue series about systemic risk.

3.1 Systemic Risk

Financial systemic risk is difficult to define and there is no consensus in the literature on a precise definition. However, proposed definitions (e.g. [37, 14, 13, 25]) all share a common notion that, when realized, systemic risk results in a severe disruption of a financial system. The outcome tends to propagate from local to system-wide shock through some amplification mechanism, also known as "contagion". Systemic risk can then be defined as the probability of such event occurring. The best recent example is the Financial Crisis of 2008 in which defaults by a small initial number of financial institutions cascaded throughout the entire US financial industry. In this case, the amplification mechanism was high interdependence of institutions' solvency due to contingent claims like default insurance. Vastly common dependence on catalytic factors like the health of the housing market also played a role [48] in this crisis. It is in general much easier to recognize systemic risk than to define it, but unfortunately recognition usually comes after the risk is realized, causing significant socioeconomic damage [37]. Policy makers therefore have a keen interest in detecting elevation of systemic risk as soon as possible in order to take preventive action.

Indeed, the predictability of systemic crises is commonly addressed in research on systemic risk. Studies that examine financial fragility at the firm level can be broadly categorized as being based on balance sheet and accounting information or market information [37]. Measures that rely on balance sheet items that have a strong relationship with the solvency of a company are only partially able to detect the risk of failure *ex ante*. One limitation of such measures is that if company management uses off-balance-sheet vehicles in operations, they potentially fail to incorporate relevant and important information for assessing the risk of insolvency. In addition, a large portion of balance sheet and accounting information is available at relatively low frequencies, e.g. on a quarterly basis. This reduces the likelihood of detection if abrupt changes in the risk take place. Finally, balance sheet information is backward-looking in the sense that it does not capture future expectations or intentions of the firm or its stakeholders.

On the other hand, measures that rely on market information are available at high frequencies and contain forward-looking information. Although market-based indicators are largely coincident with events that have been deemed of systemic importance, there is evidence that some measures can be leading indicators of elevated risk. While the difficulty of *ex ante* identification of an impending crisis is not to be underestimated, we reference [17, 40], as well as [37] and the citations therein for examples of early signals of crisis evident in market based measures.

The connection between systemic risk and correlation is rooted in a well-known stylized fact of financial returns: times of crisis are associated with increased asset return correlations. This relationship has been investigated in numerous studies, such as [6, 15, 19, 51, 55, 77]. In particular, the authors in [6] find that, conditional on the negativity of returns in the US equity market, correlations are 11.6% higher than implied by a normal distribution. This is contrasted with correlations that, when conditioned on positive returns, cannot be statistically distinguished from those implied by a normal distribution. While it has been extensively documented that correlations increase during volatile periods, a more interesting question for our purposes is whether strengthening in correlations precedes widespread shocks. That is, whether contagion is facilitated by a state of high interconnectedness that the system evolves to. In this case, the extent to which we can make inference on returns and volatility following states of high correlation is naturally of interest.

3.2 Literature Review

To measure the strength of co-movement in returns on a group of assets, we apply principal component analysis (PCA) to the sample correlation matrix and obtain the proportion of total variance explained by every PC. If a relatively small number of PCs explains a relatively large proportion of variance, this is interpreted as a state of high interconnectedness between the assets' returns. This approach has been used in the past in studies of asset return correlations, including some which analyzed correlation dynamics in the context of systemic risk [14, 20, 29, 43, 54, 60]. Since the variance explained by each PC is given by the associated eigenvalue,

this analysis can draw on what is known from random matrix theory (RMT) regarding the distribution of eigenvalues. For example, differences between the empirical density of eigenvalues and the theoretical density of eigenvalues (in purely random data) suggests that there is a common "market" factor to the returns [46]. A useful procedure is to compute the proportion of variance explained by a fixed number of PCs through time using PCA on a rolling window of sample returns. This results in a time series of the correlation measure and reveals the temporal dynamics of the interconnectedness between the assets.

For example, [43] calculate the absorption ratio for returns on the MSCI USA equity index. They find that, between January 1998 and October 2010, all 1% worst monthly drawdowns are preceded by a one-standard-deviation spike in the absorption ratio. In out-of-sample experiments, conditional on the absorption ratio exceeding a certain threshold, stocks with higher contribution to the risk of the whole system suffered statistically larger losses during the Financial Crisis of 2008 [14]. In [54], the authors study returns on US real-estate prices in 51 states and use the absorption ratio to analyze correlation dynamics. Regarding the question of whether housing bubbles can be identified in advance, the authors point to a gradual increase in the largest eigenvalue from 1993. Most recently, PCA has been applied to returns on volatilities implied from options in order to assess the systemic importance of various underlying equities [28]. Stocks whose correlation matrix of option-implied volatility returns has relatively large eigenvalues are classified as systemic and those with relatively small eigenvalues are classified as idiosyncratic. In addition, the authors apply a rolling-window PCA procedure to obtain a time series of the normalized largest eigenvalue and number of eigenvalues exceeding the Marčenko-Pastur law. Since the limits of the support of the Marčenko-Pastur law are based on white noise, eigenvalues that exceed the upper bound are considered significant in terms of the content of their information about the system's structure. The authors, find that, during high-volatility periods, the largest eigenvalue increases whereas the number of eigenvalues beyond the theoretical boundary decreases.

The most important conclusion that is common to studies using the absorption ratio to analyze the relationship between return co-movements and crises is that severe downturns are preceded by or, at least, coincide with increases in the absorption ratio. The explanation provided for this is that stronger integration facilitates the ripple effect that characterizes financial crises. However, it should be noted that these conclusions have been based on experiments that are limited to certain asset classes and time periods. Specifically, the Financial Crisis is by far the most common case study for this framework and equities are the most common asset class. Another limitation of prior studies is lack of econometric analysis of the absorption ratio (and related statistics) as an explanatory variable for returns. Instead, some important conclusions were drawn qualitatively or through backtesting without any formal tests.

In the remainder of this chapter we will attempt to fill the aforementioned gaps by analyzing correlation dynamics more systematically. To this end, we expand the universe of asset classes to include equities, bonds, currencies and CDS contracts. Additionally, apart from the Financial Crisis of 2008 (or simply the Financial Crisis), we consider events such as the Eurozone Sovereign

Debt Crisis and the Asian Financial Crisis 1997 (or simply the Asian Financial Crisis).

3.3 Data

In studying each crisis we collect data for groups of assets that played an important role in its cause and consequences. We also obtain data on at least one index to be used as a measure of distress for each crisis. A summary of asset groups and indices used is given in Table 3.7.

NYSE Financial Index

The NYSE Financial Index (NYK) includes NYSE-listed common stocks that belong to the Financial Sector according to the Industry Classification Benchmark. Components in the Index represent eighteen countries globally and several industries including banking, insurance, financial services and real estate investment. The market capitalization of NYK components represents a significant portion of the total market capitalization of the Financial Sector in the United States and globally [58]. We obtain daily logarithmic returns between January 1, 2003 and December 31, 2013 on 248 stocks that were components of the NYK as of this writing.

Credit Default Swaps on Financial Institutions

Credit Default Swaps can be thought of as insurance contracts on credit assets where one party pays premiums to another in exchange for a guarantee on the receipt of some notional amount in the event of default on the underlying asset. The premium, also called spread, on CDS contracts reflects the cost of insurance against default by some credit issuing entity. We obtain spreads for five-year senior debt CDS contracts on major global financial institutions for the period January 1, 2004 - December 31, 2013. The list of entities for this dataset is given in Table 3.1.

Financial Institutions in CDS Data		
ACE Limited	AIG Group	The Allstate Corporation
American Express	Banco Santander	Banco Bilbao Vizcaya Argentaria
Barclays PLC	BNP Paribas SA	Citigroup Inc
Commerzbank AG	Credit Agricole SA	Credit Suisse Group AG
Deutsche Bank AG	Goldman Sachs Group Inc	HSBC Holding PLC
ING Group	JP Morgan Chase	LCL SA
Lloyds Banking Group	Mitsubishi UFJ Financial Group	Morgan Stanley
Nomura Holding Inc	Royal Bank of Scotland Group	UBS AG

Table 3.1: Institutions included in dataset of Financial Sector CDS spreads.

S&P 500 Index

Daily values of the S&P 500 (SP500) index are collected for the period matching financial equity and CDS data.

Yields on European Sovereign Debt

We collect ten-year yields between January 1, 2006 and December 31, 2013 on bonds issues by the countries in Table 3.2. All of these countries had significant exposure to the Eurozone Debt Crisis due to their mutual economic ties.

Countries Represented in Bond Yield Data		
Austria	Hungary	Portugal
Belgium	Ireland	Spain
Denmark	Italy	Sweden
Finland	Netherlands	Switzerland
France	Norway	United Kingdom
Germany	Poland	

Table 3.2: Countries included in dataset of ten-year bond yields.

Credit Default Swaps on European Sovereign Debt

We collect five-year CDS spreads between January 1, 2006 and December 31, 2013 on debt issued by the countries in Table 3.3. All of these countries had significant exposure to the Eurozone Debt Crisis due to their mutual economic ties.

Countries Represented in CDS Spread Data		
Austria	Hungary	Slovakia
Belgium	Italy	Spain
France	Poland	
Germany	Portugal	

Table 3.3: Countries included in dataset of five-year CDS spreads for government-issued debt

Bloomberg European Financial Index

The Bloomberg European Financial Index (BEFINC) is a cap-weighted index of the most highly capitalized European companies that belong to the financial sector and trade on European exchanges. We obtain daily values of the index for January 1, 2006 and December 31, 2013.

Euro/US Dollar Exchange Rate

We obtain the daily spot exchange rate between the euro and US dollar (EURUSD) between January 1, 2006 and December 31, 2013. The exchange rate was efficient in responding to events throughout the crisis due to their implications on the stability, and hence the supply-demand balance, of the common currency (see Figure 3.1).

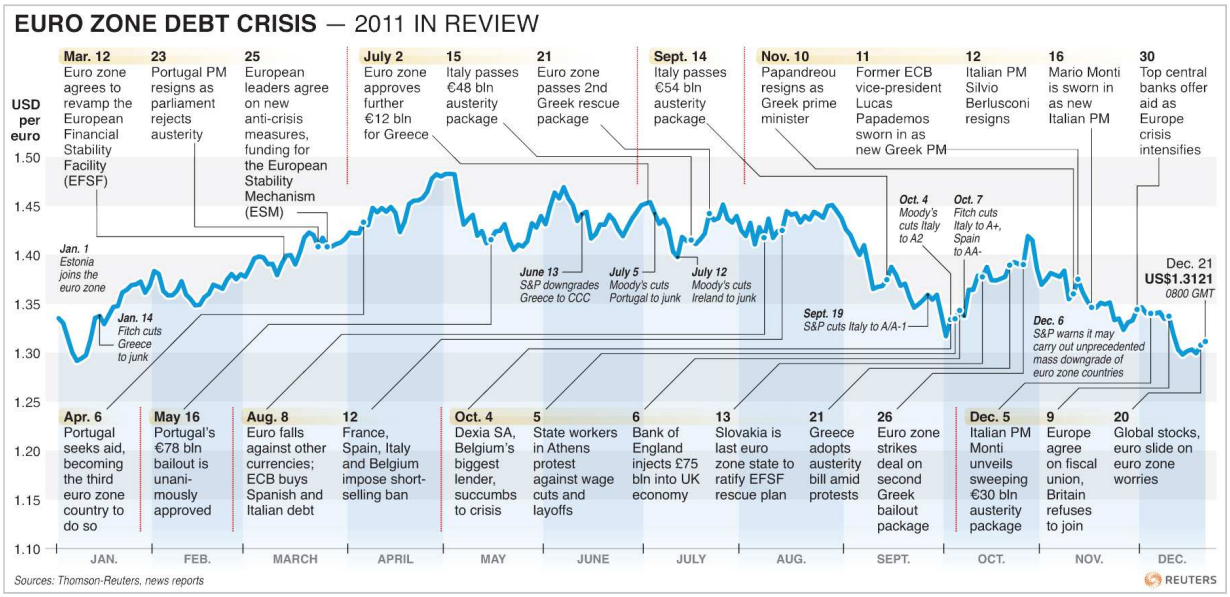

Figure 3.1: A timeline of major events related to the Eurozone Sovereign Debt Crisis and fluctuations in the EURUSD rate. Source: *Reuters*.

Currencies of Countries in Asia

We collect the spot exchange rate of the domestic currency of various countries in the region of the Asian Financial Crisis versus the U.S dollar. Included countries are listed in Table 3.4. The data is daily, for May 31, 1995 to December 31, 1999.

Countries Represented in Currency Data	
Australia (AUDUSD)	Philippines (PHPUSD)
Burundi (BNDUSD)	Singapore (SGDUSD)
Indonesia (IDRUSD)	South Korea (KRWUSD)
India (INRUSD)	Taiwan (TWDUSD)
Japan (JPYUSD)	Thailand (THBUSD)

Table 3.4: Countries (exchange rate tickers) represented in Asian currency data.

Broad Asian Equity Indices

For each country in Table 3.5, we obtain daily closing values of a major domestic free-float equity index between April 15, 1995 and December 31, 1999. The companies represented collectively in these indices account for the vast majority of economic output in the region directly affected by the Asian Financial Crisis.

Countries Represented in Asian Equity Data	
China (SHCOMP)	Philippines (PCOMP)
Hong Kong (HSI)	Singapore (SGX)
Indonesia (JCI)	South Korea (KOSPI)
Japan (NIKKEI)	Taiwan (TWSE)
Malaysia (KLCI)	Thailand (SET)

Table 3.5: Countries (equity index tickers) represented in Asian equity data.

MSCI AC Asia Pacific Index

The MSCI AC Asia Pacific Index (MXAS) captures large and mid cap representation across the 13 countries in the Asia Pacific region as listed in Table 3.6. With 989 constituents, the index covers approximately 85% of the free float-adjusted market capitalization in each country. Daily closing values are between April 15, 1995 and December 31, 1999 are obtained.

Countries Represented in Index for the Asian Financial Crisis		
Australia	China	Philippines
Hong Kong	India	Taiwan
Japan	Indonesia	Thailand
New Zealand	South Korea	
Singapore	Malaysia	

Table 3.6: Countries represented in MSCI Asia Pacific Index

3.4 Financial Crisis of 2008

Analysis of the financial sector is motivated by the importance of the Financial Crisis as a case study of systemic risk. The systemic nature of this crisis is due to links between financial institutions that were established using contingent claims such as the credit default swap (CDS). The web of contractual relationships that resulted from proliferation of claims contingent on defaults across the financial industry served as a mechanism by which losses would propagate. The interdependence in bank solvency was so profound that over 270 banks collapsed within two years of September, 2008, when Washington Mutual Inc. became the biggest bank failure on

Crisis	Assets	Indices
Financial Crisis	Equities CDSs	SP500
Eurozone Sovereign Debt Crisis	Bonds CDSs	BEFINC EURUSD
Asian Financial Crisis	Currencies Equities	MXAS

Table 3.7: Summary of asset groups and indices used in analyzing each crisis episodes. The AR is calculated for each asset group and its relationship with the level of distress is studies using the relevant indices as proxies.

record [76]. Since these links were formed and strengthened over time, it is plausible that the increased economic dependence of financial institutions resulted in a stronger co-movement in the value of the assets of these institutions.

The first dataset for the financial sector comprise daily logarithmic returns between January 1, 2003 and December 31, 2012 (2769 observations) on 248 stocks that were components of the NYK as of this writing. We estimate the correlation matrix using a rolling window of 252 trading days and perform PCA to obtain the absorption ratio. We plot the absorption ratio time series using different numbers of eigenvalues in Figure 3.2. The resulting time series are display rich dynamics and are statistically non-stationary based on ADF, PP and KPSS tests. This is particularly true for the first eigenvalue, whose relative magnitude varies dramatically, with a range from 25% to 65%. It appears that co-movement in returns has strengthened gradually in the years leading up to the Financial Crisis and remained relatively elevated since. Specifically, the average before September 1, 2008 was 34% and 53% after. Apart from being richly dynamic, $\hat{\phi}_1$ explains a relatively large proportion of the variance; it is, at its lowest, about 16 times larger than the upper support of the Marčenko-Pastur distribution with a commensurate aspect ratio parameter. On average the first PC explains approximately 44% of the variance. The first 25 PCs, a tenth of the number of variates, account for approximately 70% of the variance on average and as much as 80% at the apex of the crisis.

We also analyze 2051 observations of logarithmic returns on CDS spreads for 24 major global financial institutions. A time series of CDS spreads for some major US financial institution is plotted in Figure 3.3. As with our analysis of financial equity returns, we plot the absorption ratio time series for CDS data using a 252-day rolling window in Figure 3.4. Again, the resulting time series are richly dynamic and statistically non-stationary based on ADF, PP and KPSS tests. We also see that correlations have remained strong after the crisis in CDS spread returns. The relative magnitude of the largest eigenvalue varies approximately from 17% to 66% and much higher than the upper bound predicted by RMT. There is a dramatic spike occurring in June, 2007 that is followed by a steady increase throughout the financial crisis. On average, $\hat{\phi}_1$ explains 47% of the total variance.

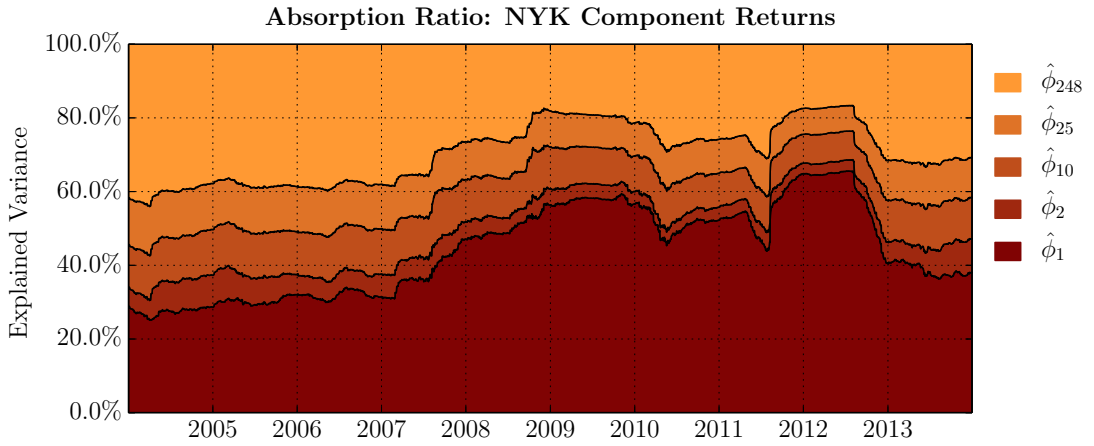

Figure 3.2: Eigenspectrum analysis for returns on components of the NYSE Financial Sector Index (NYK)

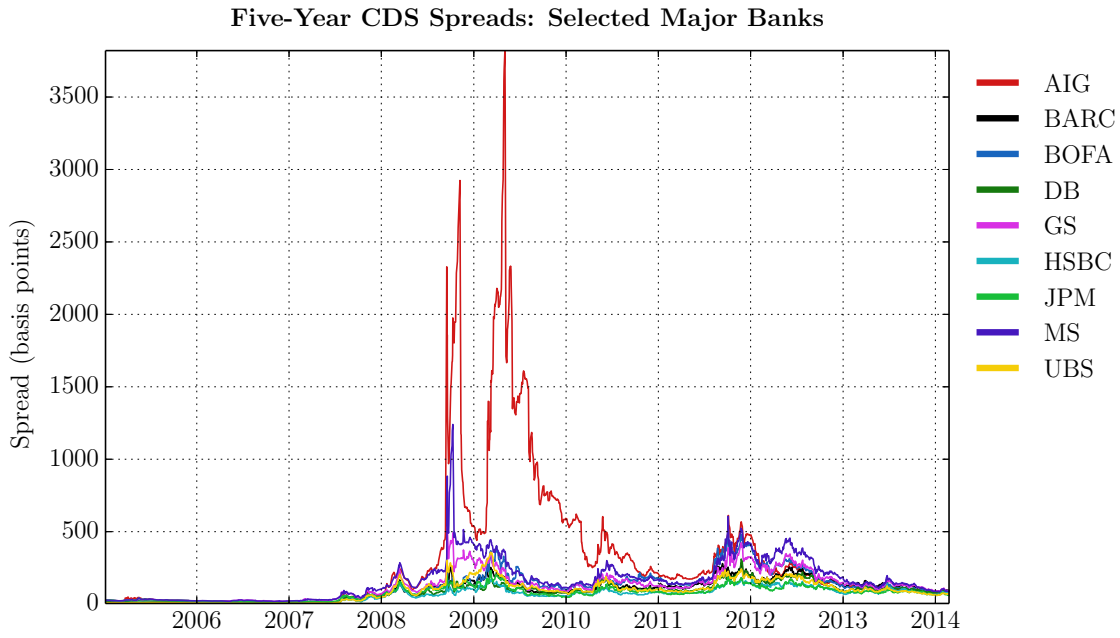

Figure 3.3: Five-year CDS spreads in basis points on selected major banks.

To illustrate the relationship between SP500 and proportion of variance explained by the first PC in our datasets, we overlay their values in Figure 3.5. Sharp increases in $\hat{\phi}_1$ are evident before critical events such as major bank failures. This suggests that a systemic risk monitoring scheme based on the level of $\hat{\phi}_1$ could potentially raise a red flag before the crisis reached its height. However, if there is any predictive component to the information content of $\hat{\phi}_1$, it is certainly mixed with some coincident response to shocks. For example, we observe a dramatic spike in equity correlations in the summer of 2011, which coincides with the credit rating downgrade of

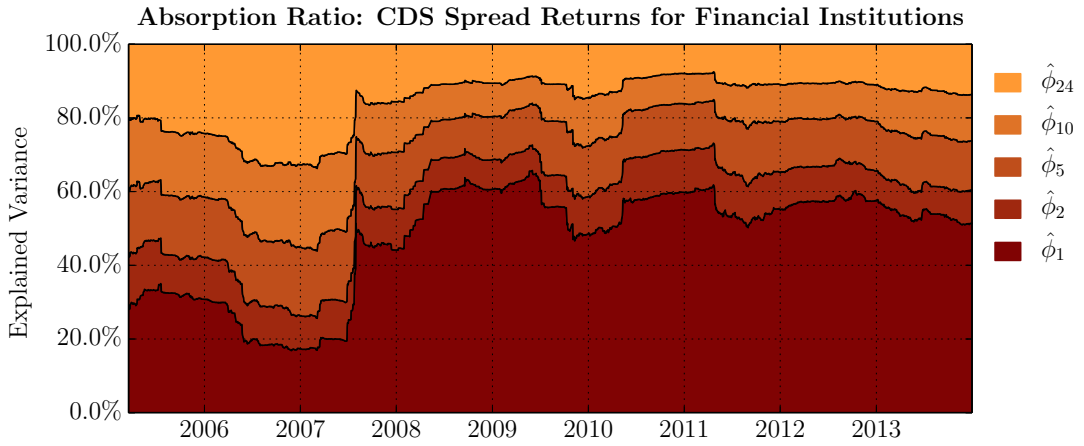

Figure 3.4: Eigenspectrum analysis for returns on spreads of financial institution CDSs

US sovereign debt. However, there is no apparently significant change in the correlation structure prior to this event.

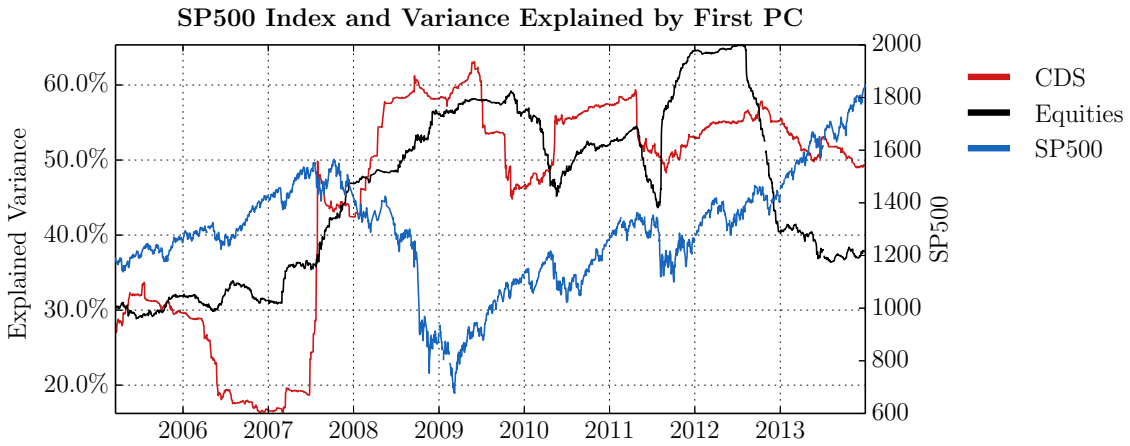

Figure 3.5: Proportion of variance explained by first PC in CDS and NYK component (equity) data, compared with the level of SP500.

Our observations thus far are consistent with existing work in the literature. Periods of turmoil in equity markets appear to have been preceded by or coincide with increases in correlation of returns.

3.5 Eurozone Sovereign Debt

Compared to previous crises, the Financial Crisis took place in a system featuring a higher level of globalization in business and finance. Increased financial globalization is driven by regulatory and technological factors that contribute to, among other things, higher cross-border capital flows,

sharing of information or best practices among financial institutions, centralized exchanges, standardized contingent claims and other mechanisms simplifying foreign participation in domestic financing activity. In an environment of this nature, economic cycles in different economic regions are expected to be able align. In addition, formation of asset bubbles due to, for example, a rising level of credit, can occur simultaneously and spread across multiple regions. It was indeed the case that economic recession and collapse of real-estate prices in the late 2000s occurred on a multinational scale over a relatively short time period. In several eurozone countries, these events brought to light a troubling fiscal account, resulting from a combination of profligate government spending and debt nationalized in efforts to support massive private sector losses associated with the real-estate crash. In particular, a common pitfall for these countries was excessive borrowing when, by virtue of the monetary union, their creditworthiness was considered similar to stronger eurozone countries. Overestimated credit quality meant overpriced debt, and low interest rates. The cost of borrowing began to rise with perceived risk in Greece in late 2009. The loss of confidence quickly spread to other countries including Ireland, Italy, Portugal and Spain, all of which saw their cost of borrowing and default insurance rise dramatically.

The eurozone sovereign debt crisis is an example of a systemic shock in fixed income assets. It is similar to the Financial Crisis in that it was characterized by financial contagion. While the heart of the crisis involved certain members of the currency union, known as peripheral countries, financial inter-linkages meant that countries that are not necessarily in the eurozone would be affected, by virtue of their membership in the European Union, for example. Thus the analysis in extant literature and our work so far is consistent with the setting of this crisis. We extend it here by studying the relationship between correlation and contagion in the European sovereign credit market.

We focus on the information contained in returns on sovereign bond yields and CDS spreads of countries in the EU[1]. The first dataset is comprised of daily logarithmic returns on ten-year yields for government debt issued by the countries listed in Table 3.2. The second dataset consists of daily logarithmic returns on five-year CDS spreads on sovereign debt issued by the countries in Table 3.3.

Note that the variates in the system ought be temporally homogeneous in order to preserve comparability and the power of inferential analysis. A missing value in one variate would require filling or else the time point would need to be discarded altogether. Therefore we have excluded certain countries from the analysis due to limited data availability. For example, Greece stands out as the country that had endured the most significant rise in its cost of borrowing, as well as CDS spreads reflecting expectations of near certain default. Greece indeed practically defaulted on its debt as it had undergone restructuring, with private bondholders accepting deep haircuts, and a downgrade of its debt to 'Restricted Default' rating [71]. As a result data on bond yields

[1]The vast majority of the countries were *de jure* EU members at the time of writing, though we also study countries who have adopted provisions in order to participate in the EU single market without membership.

and CDS spreads was not available after March 9, 2012 and September 16, 2011, respectively.[2] Comparing results using samples with and without data on Greek debt for available time periods we find that, despite Greece's central role in the crisis, the results are not materially different. For this reason exclude data on Greek debt from our analysis.

The cost of debt and its insurance would reach historically high levels for multiple countries, particularly Greece, Ireland, Italy, Portugal and Spain. A time series plot of yields for selected eurozone countries is presented Figure 3.6. A similar plot of CDS spreads is presented in Figure 3.7. It can be seen both in yield and CDS spread data that concerns regarding sovereign creditworthiness in the eurozone started to mount in 2009.

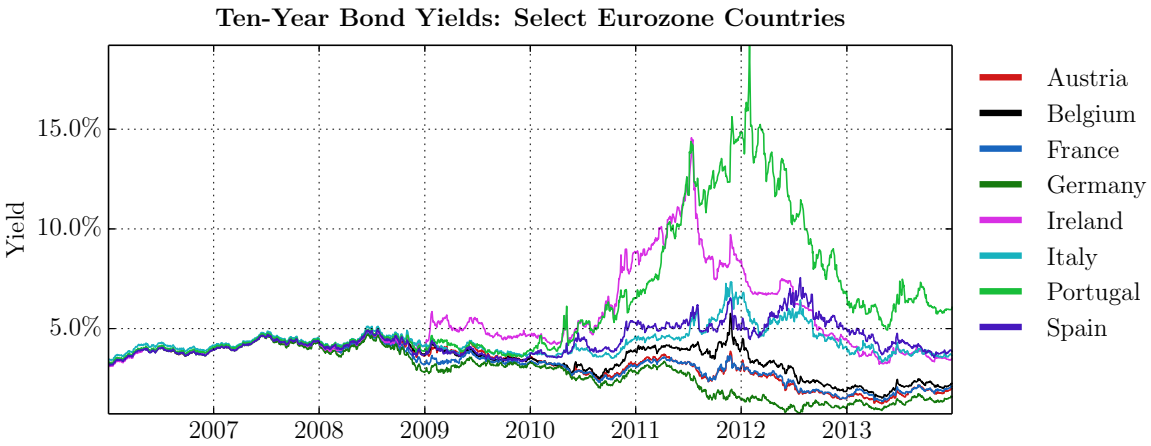

Figure 3.6: Ten-year yields on sovereign debt issued by selected eurozone member states.

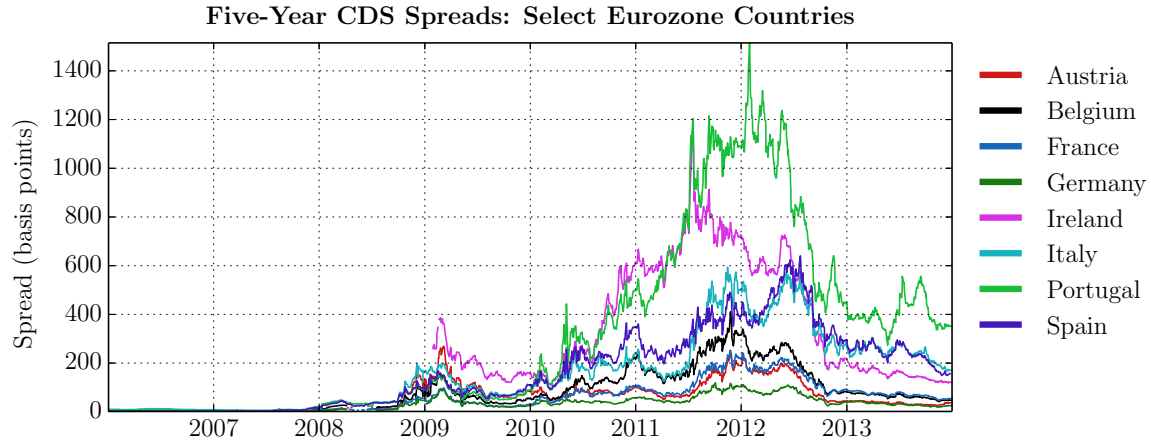

Figure 3.7: Five-year CDS spreads in basis points on selected eurozone member states.

In calculating the absorption ratio, we keep the size of our rolling estimation window at 252 trading days. Figure 3.8 plots the temporal evolution of the absorption ratio for bonds in this system. The correlation structure is non-stationary by casual observation and based on ADF,

[2] Greek bond yield data is missing for approximately one year.

PP and KPSS tests. Otherwise, we observe that the correlation dynamics in this case are very different from our results for the Financial Crisis. There is a notable decline in the absorption ratio of a small number of eigenvectors through the crisis. Returns on sovereign bond yields in Europe were highly correlated before the Financial Crisis, with $\hat{\phi}_1$ explaining approximately 80% of the variance. Interestingly, near the end of 2008, when the Financial Crisis was at its peak, the strength of co-movement began to decrease steadily. There was also a sharp shift in the structure at the start of May 2010, when the first PC decreased in dominance compared with the rest. The aspect ratio is 0.0675 and the Marčenko-Pastur density with this parameter has an upper support of 1.58 so that the limiting theoretical (normalized) maximum of the largest eigenvalue, assuming white noise data, is approximately 9.3%, or about 7 times smaller than what we observe on average.

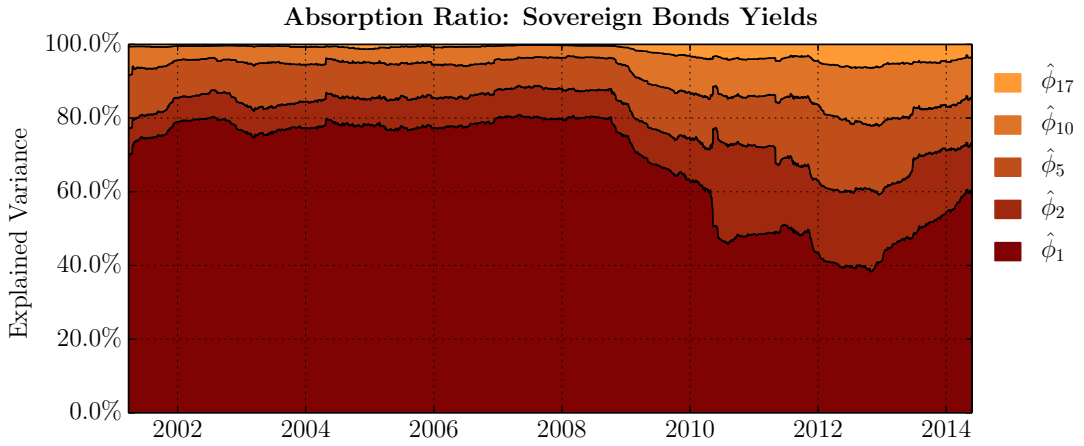

Figure 3.8: Eigenspectrum analysis for returns on eurozone sovereign bond yields.

It can also be seen that $\hat{\phi}_1$ began to increase through 2013, as yields and CDS spreads have come down in response to improving conditions. Our results suggest that the co-movement in bond yields weakened as a result of distress. Said otherwise, bond yield fluctuations have diverged as bonds of certain countries became an alternate, safer, source of yield vis-à-vis those experiencing fiscal distress. For example, German, French and Austrian bond yields were decreasing while Portuguese, Italian, Irish and Spanish yields were rising for certain periods (see Figure 3.6). This is interesting because decreasing yields should reflect an improvement in the issuing entity's creditworthiness. Analysis of the creditworthiness of countries is beyond the scope of this work. However, between increasing and decreasing risk of default among so-called"core" countries, we believe that the more plausible option is increasing risk, due not least to uncertainties regarding the stability of the euro currency, which is common to both core and peripheral countries. At least in theory, the mere exposure of core country governments and banks to distressed debt would have put, *ceteris paribus*, upward pressure on core country yields as well. Thus, the divergence in yield fluctuations would be in contrast to the fundamental risk-return relationship and would serve to highlight the behavioural aspects in asset pricing.

Next, we look at the absorption ratio for CDS spreads, plotted in Figure 3.9. The correlation dynamics for sovereign debt CDS spread fluctuations resemble their analogue for debt issued by our sample of financial institution in Section 3.4. We see a gradual increase in the absorption ratio starting in the latter half of 2008 and continuing through 2009. In this shift $\hat{\phi}_1$ rises from about 23% to near 60%. This implies a significant increase in the strength of co-movement between CDS spreads. We also find that co-movement has begun a gradual yet significant weakening starting in late 2012. Based on a comparison with results from RMT, as before there is evidence of non-random structure since the largest eigenvalue is, on average, about 4 times greater than the theoretical upper bound.

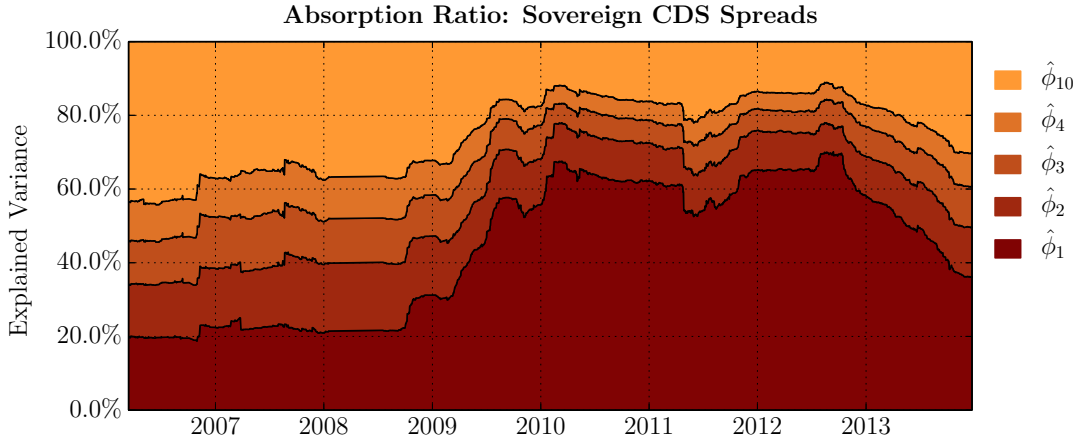

Figure 3.9: Eigenspectrum analysis for returns on eurozone sovereign CDS spreads.

It is evident that the correlation dynamics of bond yields and CDS data were different during the eurozone debt crisis. Our results suggest an inverse relationship where bond yield fluctuations diverged during market distress and converged in times of relative calm. On the other hand CDS spreads returns converged in times of distress and diverged otherwise, in agreement with what we have observed when analyzing correlations during the Financial Crisis. These relationships are visualized in Figure 3.10, where the largest eigenvalue of each asset class is compared with BEFINC, and likewise in Figure 3.11 for the EURUSD rate.

To the extent possible, we aim to study the Eurozone Debt Crisis in isolation, despite its clear connection with the events of the Financial Crisis. In other words, our interest is in the correlation dynamics of a basket of assets that were central to the sovereign debt crisis and their relationship with a broader measure of the system's economic stability. Before recognition of fiscal trouble in the eurozone, such measures reflect distress from the Financial Crisis that is independent from the effects of what is now deemed as a loss of confidence in the ability of multiple countries to make good on loans. Therefore in studying the connection between correlations and market turmoil as per a broad, observable indicator, our period of study for this crisis begins on March 9, 2009[3].

[3]Major equity indices in the US have reached their low on this date and it has been used as a starting point for post crisis analysis in reference to the Financial Crisis. [23]

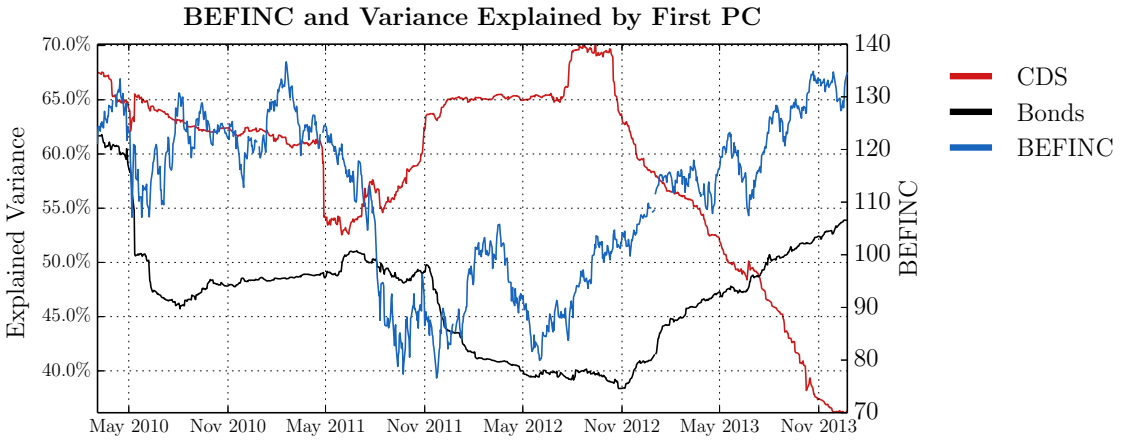

Figure 3.10: Proportion of variance explained by first PC in CDS and bond yield data, compared with the level of BEFINC.

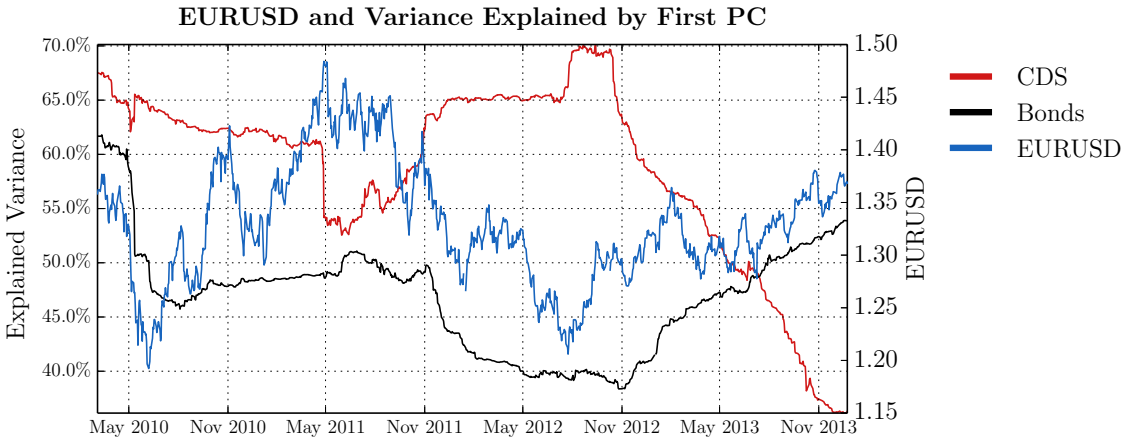

Figure 3.11: Proportion of variance explained by first PC in CDS and bond yield data, compared with the level of EURUSD.

3.6 Asian Financial Crisis

In 1997 the global financial system experienced the ripples of a systemic shock to the economies of a few countries in Asia. Named after its origin, the Asian Financial Crisis is another case study of financial distress that was contagious due to strong economic integration between the affected entities. It is yet another example of how the financial difficulties of one or few components in a system can affect it on a much wider scope. The distinguishing aspect of this crisis from events discussed so far is that currencies played a central role in its development. To the best of our knowledge, our examination of correlation dynamics of currencies in times of distress is novel.

As with many other crises, the exact cause and starting point of this crisis is debatable. One proposed view is that the main source of initial financial turmoil was sudden shifts in market expectations and confidence. And while economic fundamentals somewhat deteriorated in the

affected countries, this should not be identified as a cause as much as panic on the part of investors. An alternate view is that the crisis was triggered by fundamental imbalances and pushed by behavioural factors to be more adverse than would be justified fundamentally [22]. In either case the hotbed of turmoil is considered to be the currency market by most participants [33] so the macroeconomic background of the crisis is deemed important.

In the decade before the crisis several Asian countries experienced unprecedented economic growth. Although there were important differences between the individual countries, several common elements lead to stronger association between the economic welfare of these countries. These were likely both fundamental imbalances – of which the crisis reflected reversion towards equilibrium – and circumstantial elements which promoted concerted investor sentiment about these countries as a group more so than individually.

On the fundamental side, East Asian economies were commonly export-driven due to a combination of inexpensive and relatively well-educated labour and falling barriers to international growth. The wealth created by strong exports contributed to an investment boom, much of which was financed with debt denominated in foreign currency, particularly the US dollar. These investments were primarily made in infrastructure, industrial capacity and commercial real estate, which required large quantities of foreign goods. All the while, political pressures to maintain high rates of economic growth had led to a tendency by government to guarantee private projects. Thus many East Asian countries saw growing deficits and excess capacity [22]. Importantly, much of the borrowed capital had been in US dollars because local currencies were pegged to the dollar and debt in domestic currency carried a higher interest rate. However, maintaining the peg of a domestic currency to the dollar required large quantities of foreign reserves to support it.

In late 1996 and early 1997 Thailand's currency, the baht, was subject to speculative attacks from traders. Short sellers had begun to recognize the weakness of the baht in terms of its relatively low demand compared to the US dollar, which was required to service debt, and a diminishing ability of the government to defend it with (still depleting) foreign currency reserves. In July 1997, when the Thai government was no longer able to defend the baht, it was allowed to float freely. The result was an immediate sharp decrease in its value. The baht continued to depreciate through the year, making the debt burden of Thai companies increasingly difficult to manage to the point that many of them were forced into bankruptcy. At the same time these events were coupled with a significant decline in the Stock Exchange of Thailand.

Panic spread to neighbouring countries as various Asian currencies experienced similar speculative attacks. Within two months, depleting foreign exchange reserves forced Malaysia, Singapore and Indonesia to drop the peg of their currencies to the US dollar. In each case, the move was met with a sharp devaluation of the local currency and equities. The shock would significantly affect South Korea, Japan, Taiwan, Philippines, Laos. This can be seen in Figure 3.12 and Figure 3.13, where we plot the returns currencies and equity indices of several Asian countries. Other countries were affected to a lesser extent, such as China, which saw reduced growth rates following the crisis.

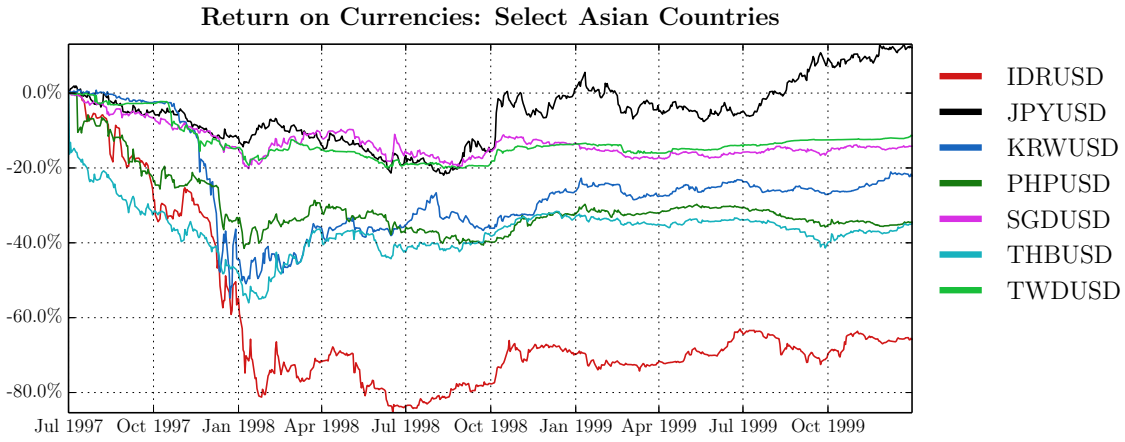

Figure 3.12: Returns on currencies of Asian countries affected by the Asian Financial Crisis

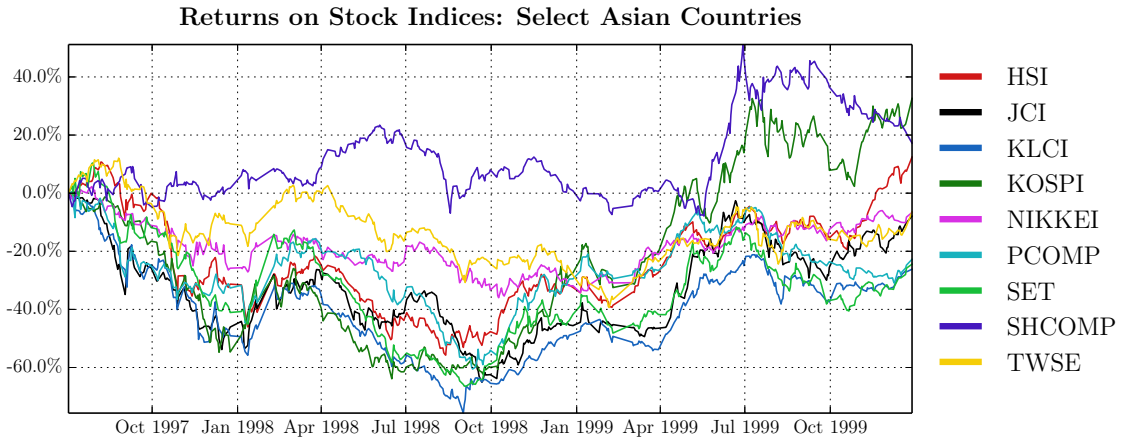

Figure 3.13: Returns on equities of Asian countries affected by the Asian Financial Crisis

The developments in this particular crisis make it a compelling subject of investigation in our study due to the spectacular contagion effect. We again seek to gain insight into the correlation dynamics of tradable assets and potential relationships with returns in times of distress. The assets that we investigate are a basket of currencies of Asian countries and a basket of major stock indices. It is important to note that for the period that a given currency is pegged, the price discovery process is critically altered by government intervention. This affects any conclusions drawn about the dynamics of the AR prior to the crisis. However, the returns of pegged currencies are not identically zero as can be seen in Figure 3.14. The currencies fluctuate within a narrow band under normal conditions. During an unusual trading activity, such as speculative attacks, fixed exchange rate bands can be breached. Therefore there is still useful information content in currency returns even under a fixed exchange-rate policy.

We maintain the size of our rolling estimation window for the AR at 252 trading days. Figure

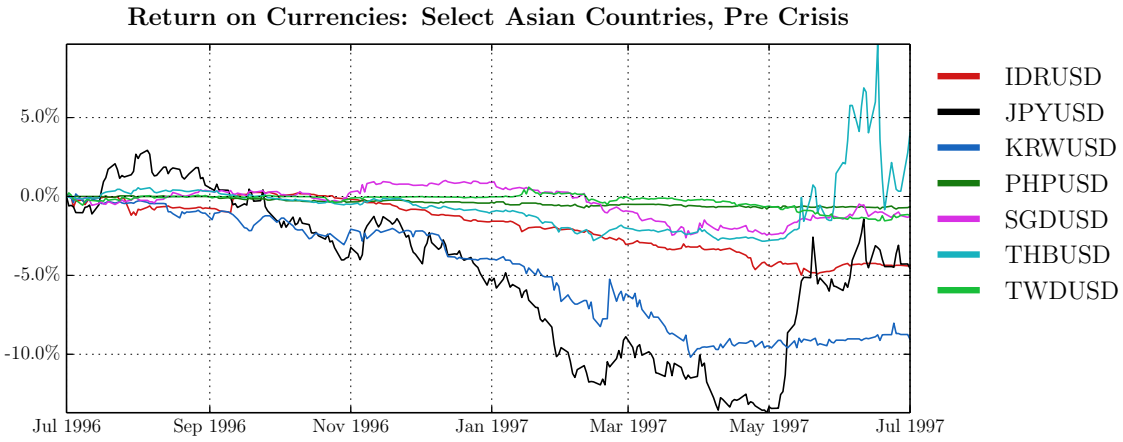

Figure 3.14: Returns on currencies of select Asian countries before local currency pegs to the US dollar were dropped by affected countries.

3.15 plots its the temporal evolution for returns on Asian currencies versus the US dollar. The correlation structure appears to be non-stationary and stationarity is rejected based on ADF, PP and KPSS tests. We observe a strengthening in correlation beginning in mid to late 1997. Similar features are present in equity data, which we use to generate the AR series In Figure 3.16. The aspect ratio of the correlation matrices for both asset baskets is $y = \frac{10}{252}$. Nevertheless the distribution of eigenvalues differs, with equity returns being more tightly coupled as $\hat{\phi}_k$ is higher in equities for each k. In particular, $\hat{\phi}_1$ is 31% for equities and 21% higher on average over the same period. With the aspect ratio at hand, the upper support of the Marčenko-Pastur density is 14%.

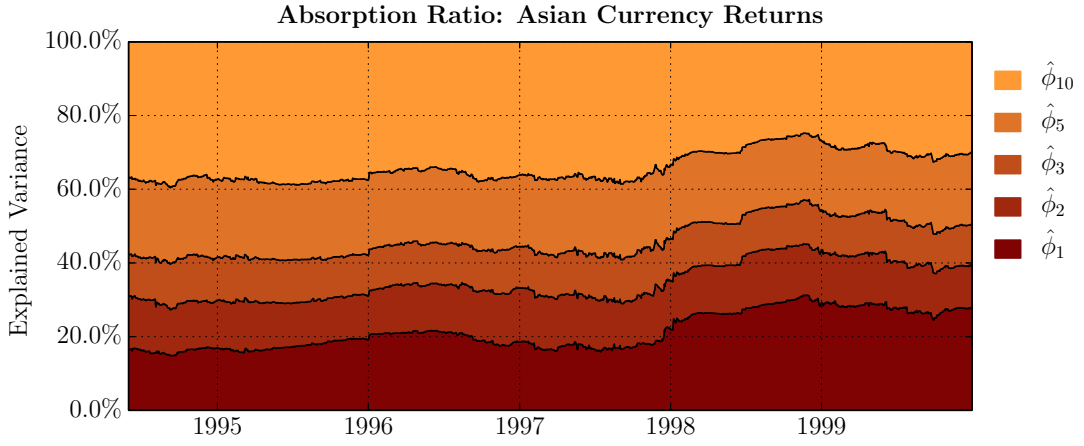

Figure 3.15: Eigenspectrum analysis for returns on currencies of countries in Asia that were significantly affected by the Asian Financial Crisis.

Both equities and currencies appear to have increasingly correlated returns in times of crisis. An inverse relationship between the AR and the MXAS index can be seen in Figure 3.17. For

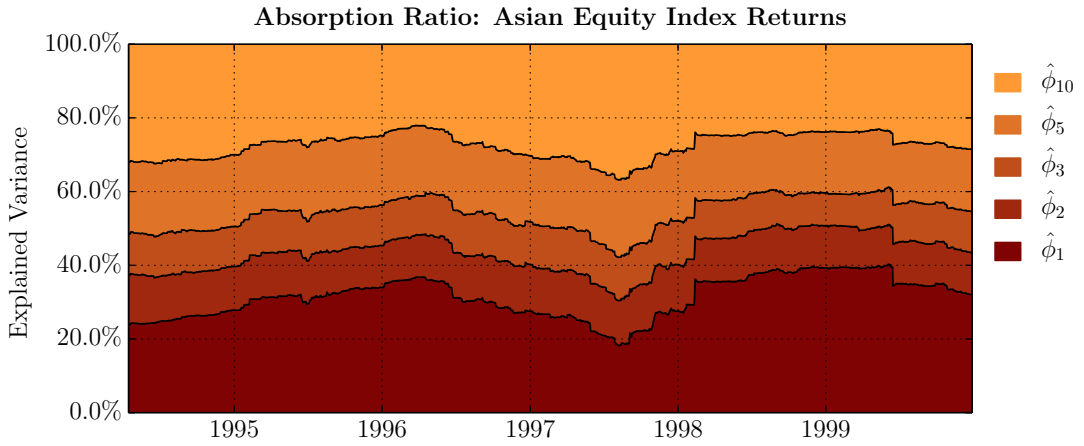

Figure 3.16: Eigenspectrum analysis for returns on major stock indices of countries in Asia that were significantly affected by the Asian Financial Crisis.

equities, this supports our findings from Section 3.4. Currencies have been unexplored in this type of framework before. Apart from the directional relationship there is the important question of whether correlations tend to strengthen before crisis episodes and thereby foster the type of contagion that characterizes systemic events. Figure 3.17 provides evidence to the contrary. Returns on equities and currencies in the region had been decoupling from mid 1996 until the onset of the crisis.

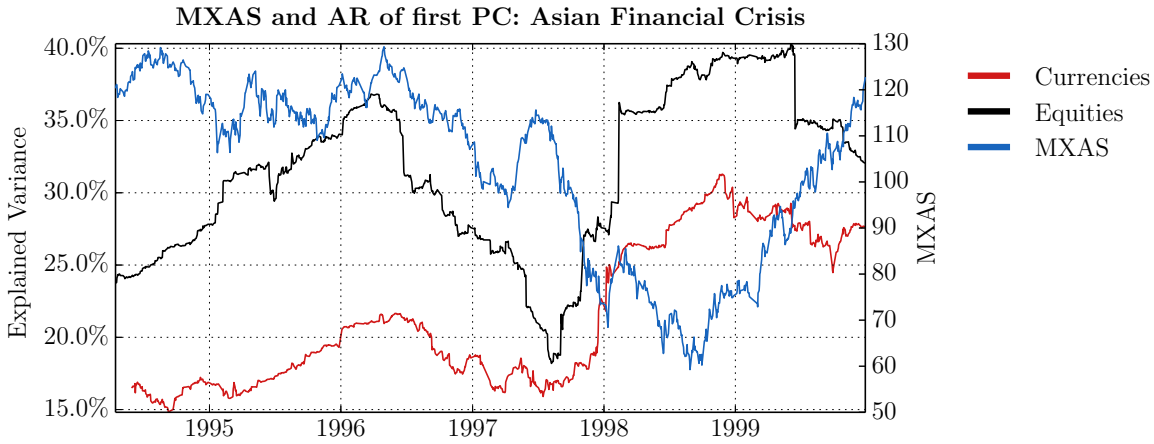

Figure 3.17: Proportion of variance explained by first PC in currency and equity yield data, compared with the level of MXAS.

Plotting a time series of the AR provides a visual summary of the strength of co-movement in the data variates. In this section we found some evidence for a relationship between the AR and the level of distress in the system as measured by a broad index during various crises. We found that returns on stocks and CDS spreads of financial institutions appeared to have an inverse relationship with the NYSE Financial Sector Index during the Financial Crisis. Furthermore,

correlations in these assets had been increasing before the crisis attained its peak. Returns on CDS spreads of European sovereign debt had a similar behaviour around the Eurozone Sovereign Debt Crisis. However, returns on yields of bonds issued by these countries had become less correlated. Finally, higher correlation in stock and currency returns around the Asian Financial Crisis appears to be associated with higher levels of distress.

3.7 Regression Analysis

In Section 3.2 it was pointed out that the existing literature broadly agrees with the conclusion that financial distress episodes are associated with increasing correlations at least contemporaneously (and possibly *ex-ante*), and such conclusions are based on analysis of the most recent major financial crisis, in 2008. While some of our results support this, taken in total they imply that there exists a more general relationship. Specifically, we have qualitative evidence to suggest that crises are linked to a breakdown in correlation structure – one which does not necessarily result in a strengthening of co-movement but possibly in a decoupling of related assets. The divergence of sovereign bond yields in the Eurozone Sovereign Debt Crisis is a clear example of the latter. In this section we make a formal analysis of the relationship between the AR and broad market conditions using a linear regression technique.

Our objective in this sub-section is to determine whether there is a statistically significant association between changes in the AR and the returns or volatility of a broad index. In this setting the index would act as a proxy measure of financial distress. The relationship between changes in the AR and drawdowns in the MSCI USA Index was explored in [43]. The authors find that all of the 1% worst monthly drawdowns are preceded by a one-standard-deviation spike in the AR[4]. However, no econometric test for a relationship is conducted. We will fill this gap with linear regression models based on our hypothesized relationships. We are also interested in investigating whether any potentially significant relationships are leading, coincidental or lagging in nature. It would be considered leading if shifts in the AR are better at predicting returns or volatilities than vice versa. If the opposite is true, then the relationship would be lagging. To this end we test for Granger causality and also examine significance of coefficients in linear regression models with leading and lagging terms.

3.7.1 Granger Causality

The causal link between two random variables X and Y can be defined in terms of predictability in linear regression models as follows.

Definition 12 (Granger-Causality)**.** Let $\{X(t)\}$ and $\{Y(t)\}$ be two time series. Further, let $\mathcal{I}_X(t) = \big[X(t), X(t-1)\dots\big]$ be the information about X at time t, and similarly

[4]The converse was not asserted to be true. In other words, not all one-standard-deviation spikes in the AR were followed by a drawdown in the 99th percentile.

$\mathcal{I}_Y(t) = \big[Y(t), Y(t-1)\dots\big]$. We say that $X(t)$ *does not Granger-cause* $Y(t)$ if and only if, conditional on $\mathcal{I}_Y(t-1)$, $Y(t)$ is independent of $\mathcal{I}_X(t-1)$. Otherwise, we say that $X(t)$ *Granger-causes* $Y(t)$.

For our purposes it is of interest to examine whether there exists such a causal relationship between changes in the AR and the returns or volatility of an index series which serves as a proxy for the financial condition of a system in distress. Consider a bi-variate VAR model [36]

$$r_Y(t) = \alpha_0 + \sum_{i=1}^{k} \alpha_i r_Y(t-i) + \sum_{i=1}^{k} \beta_i \Delta\phi_1(t-i) + \epsilon(t) \tag{3.1}$$

$$\Delta\phi_1(t) = \gamma_0 + \sum_{i=1}^{k} \gamma_i \Delta\phi_1(t-i) + \sum_{i=1}^{k} \delta_i r_Y(t-i) + \eta(t) \tag{3.2}$$

where $\Delta\phi_1(t)$ is change in the AR of the dominant PC, $r_Y(t)$ is the return on index Y, k is the maximum number of lagged observations included in the model, $\epsilon(t)$ and $\eta(t)$ are taken to be serially uncorrelated white noise variables. If the variance of ϵ is reduced by the inclusion of $\Delta\phi_1$ terms in Equation 3.1 then $\Delta\phi_1$ Granger-causes r_Y. Similarly, if the variance of η is reduced by the inclusion of r_Y terms in Equation 3.2 then r_Y Granger-causes $\Delta\phi_1$. A null hypothesis that $\Delta\phi_1$ does not Granger-cause r_Y is rejected if any of the coefficients of r_Y are jointly significantly different from zero with those of $\Delta\phi_1$ according to an F-test, and vice versa. A variable that Granger-causes another is considered to be a leading source of information for predicting it. One may find that neither variable Granger-causes the other or that the variables Granger-cause each other. In the former case there may be no relationship between them whereas in the latter there may be a feedback relationship.

We fit the model using ordinary least squares and perform these tests separately for returns, r_Y, and realized volatility, σ_Y. In addition, we compute changes in AR, returns and realized volatilities over different time intervals and repeat the analysis for robustness. Specifically, we use daily, weekly, monthly and quarterly changes. The maximum lag in the fitting of each VAR model is determined using the AIC criterion. Components underlying the AR and the corresponding index depend on the crisis episode and are commensurate with Section 3.4 - Section 3.6. These are also summarized in Table 3.7.

Results of Granger causality tests are presented in Table 3.8 - Table 3.11. Each table corresponds to results with respect to a specific market index, Y. Around the Financial Crisis, changes in correlation between stocks that were components of the NYK are found to Granger-cause the realized volatility of the broader market. However there is insufficient evidence to reject the no-causality hypothesis based on CDS data. In the case of index returns we find evidence to support the reverse direction of causality both in equities and CDSs. For the Eurozone Sovereign Debt Crisis we find evidence of causality between changes in AR and volatility of both the BEFINC and

EURUSD, consistently across bond yields and CDS spreads. The direction of causality suggests that changes correlation strength occurred in response to changes in market conditions, although in CDS data there appears to be a feedback effect if the EURUSD was to be taken as an index. There is insufficient evidence to make similar conclusions with respect to index returns except in a few isolated cases. In our period of study containing the Asian Financial Crisis, we cannot reject that changes in correlation of currency or equity returns did not have a causal relationship with either the returns or volatility of the MXAS, except in isolated cases. Overall these results support the idea that there is a cause-and-effect relationship, although the direction is not consistent. Conclusions from these tests are made about causality in a predictive sense. Thus, failure to reject the no-causality null hypothesis, even in both directions, does not imply that there is no association between the variables. For example, the test is not designed to account for contemporaneous relationships or common causality with a third variable. Nevertheless we conclude that within our battery of tests there are noteworthy patterns in the interplay between changes in the AR and market turmoil, in terms of volatility of returns more so than returns.

3.7.2 Dimson Regression

Our second approach to analyze the lead-lag effects between correlation and distress is based on Dimson regression [27]. The model for return of index Y is given by

$$r_Y(t) = \alpha_0 + \sum_{i=-k}^{k} \beta_i \Delta\phi_1(t-i) + \epsilon(t), \tag{3.3}$$

where ϵ is a constant-variance Gaussian error term. For volatility we simply replace $r_Y(t)$ by $\sigma_Y(t)$ in an independent regression. The coefficients β_i with negative subscripts are lag coefficients and those with positive subscripts are lead coefficients. We can test the significance of each coefficient using a t-test on the null hypothesis that it is equal to zero. If any lag coefficients are found to be significantly different from zero then it can be inferred that returns or volatility lag changes in the AR. On the other hand, rejecting the null hypothesis for any of the lead coefficients would support the notion that changes in the AR lag returns or volatility. There is also one contemporaneous coefficient, β_0, the significance of which would imply that changes in asset correlations occur simultaneously with returns or volatility in the system as captured by the index. Finally if, together with a rejection of the null hypothesis with the contemporaneous coefficient, we find that both lead and lag coefficients are different from zero then the relationship is considered informationally efficient.

We fit the model using ordinary least squares. In conducting the t-tests we use White's heteroskedasticity-consistent standard errors [85]. We select $k = 1$ but, as in Section 3.7.1, we compute changes in AR, returns and realized volatilities on daily, weekly, monthly and quarterly basis. Thus we study the effects of a one-period lead or lag over different observation windows. Components underlying the AR and the corresponding index depend on the crisis episode and

Index: SP500	$\Delta\hat{\phi}_1(t) \to Y$		$\Delta\hat{\phi}_1(t) \leftarrow Y$	
	F-stat	p-value	F-stat	p-value
Equities				
Return (Daily)	0.34	0.79	7.96	0.00***
Return (Weekly)	0.44	0.72	3.71	0.01**
Return (Monthly)	0.03	0.86	7.24	0.00***
Return (Quarterly)	0.07	0.78	0.99	0.32
Volatility (Daily)	4.01	0.00***	96.30	0.00***
Volatility (Weekly)	2.55	0.05*	1.14	0.32
Volatility (Monthly)	2.36	0.09*	1.19	0.30
Volatility (Quarterly)	0.55	0.46	0.00	0.92
CDSs				
Return (Daily)	2.42	0.08*	1.13	0.32
Return (Weekly)	0.15	0.69	0.52	0.46
Return (Monthly)	0.98	0.40	4.50	0.00***
Return (Quarterly)	0.01	0.91	9.18	0.00***
Volatility (Daily)	0.08	0.96	1.51	0.20
Volatility (Weekly)	0.52	0.66	0.21	0.88
Volatility (Monthly)	0.39	0.67	1.76	0.17
Volatility (Quarterly)	1.86	0.16	1.64	0.20

* statistically significant at 10%

** statistically significant at 5%

*** statistically significant at 1%

Table 3.8: Granger-causality test results for the Financial Crisis. Each row represents a test of causality between either returns or realized volatility of SP500 and changes in the AR as derived from equity or CDS data, over various time frequencies. The F-stat and p-value found in each test are given for both directions of causality, with $\Delta\hat{\phi}_1(t) \to Y$ denoting that changes in ϕ_1 Granger-cause the index variable. Conversely $\Delta\hat{\phi}_1(t) \leftarrow Y$ denotes that the index variable Granger-causes changes ϕ_1.

Index: BEFINC	$\Delta\hat{\phi}_1(t) \rightarrow Y$		$\Delta\hat{\phi}_1(t) \leftarrow Y$	
	F-stat	p-value	F-stat	p-value
Bonds				
Return (Daily)	1.67	0.17	4.02	0.00***
Return (Weekly)	0.86	0.35	3.95	0.04**
Return (Monthly)	0.12	0.72	0.01	0.90
Return (Quarterly)	0.61	0.44	0.35	0.55
Volatility (Daily)	2.89	0.03**	8.31	0.00***
Volatility (Weekly)	2.75	0.06*	3.08	0.04**
Volatility (Monthly)	1.04	0.37	6.34	0.00***
Volatility (Quarterly)	0.14	0.70	0.33	0.56
CDSs				
Return (Daily)	0.59	0.54	0.11	0.89
Return (Weekly)	0.15	0.69	0.00	0.99
Return (Monthly)	0.00	0.98	1.50	0.22
Return (Quarterly)	0.63	0.60	7.23	0.00***
Volatility (Daily)	0.29	0.82	6.55	0.00***
Volatility (Weekly)	1.21	0.29	5.03	0.00***
Volatility (Monthly)	0.79	0.37	8.17	0.00***
Volatility (Quarterly)	0.50	0.48	2.01	0.16

* statistically significant at 10%

** statistically significant at 5%

*** statistically significant at 1%

Table 3.9: Granger-causality test results for the Eurozone Sovereign Debt Crisis. Each row represents a test of causality between either returns or realized volatility of BEFINC and changes in the AR as derived from bond or CDS data, over various time frequencies. The F-stat and p-value found in each test are given for both directions of causality, with $\Delta\hat{\phi}_1(t) \rightarrow Y$ denoting that changes in ϕ_1 Granger-cause the index variable. Conversely $\Delta\hat{\phi}_1(t) \leftarrow Y$ denotes that the index variable Granger-causes changes ϕ_1.

Index: EURUSD	$\Delta\hat{\phi}_1(t) \rightarrow Y$		$\Delta\hat{\phi}_1(t) \leftarrow Y$	
	F-stat	p-value	F-stat	p-value
Bonds				
Return (Daily)	0.82	0.48	1.11	0.34
Return (Weekly)	1.52	0.21	6.56	0.01**
Return (Monthly)	2.75	0.06*	2.33	0.10
Return (Quarterly)	0.95	0.33	0.40	0.52
Volatility (Daily)	2.19	0.08*	1.36	0.25
Volatility (Weekly)	1.18	0.31	2.70	0.04**
Volatility (Monthly)	0.92	0.39	2.06	0.13
Volatility (Quarterly)	0.04	0.95	5.55	0.01**
CDSs				
Return (Daily)	2.70	0.04**	0.74	0.52
Return (Weekly)	1.43	0.23	7.30	0.00***
Return (Monthly)	1.24	0.29	1.81	0.16
Return (Quarterly)	0.00	0.93	0.72	0.40
Volatility (Daily)	2.71	0.04**	3.37	0.01**
Volatility (Weekly)	2.12	0.09*	2.88	0.03**
Volatility (Monthly)	5.08	0.02**	6.80	0.01**
Volatility (Quarterly)	0.00	0.92	2.13	0.15

* statistically significant at 10%

** statistically significant at 5%

*** statistically significant at 1%

Table 3.10: Granger-causality test results for the Eurozone Sovereign Debt Crisis. Each row represents a test of causality between either returns or realized volatility of EURUSD and changes in the AR as derived from bond or CDS data, over various time frequencies. The F-stat and p-value found in each test are given for both directions of causality, with $\Delta\hat{\phi}_1(t) \rightarrow Y$ denoting that changes in ϕ_1 Granger-cause the index variable. Conversely $\Delta\hat{\phi}_1(t) \leftarrow Y$ denotes that the index variable Granger-causes changes ϕ_1.

Index: MXAS	$\Delta\hat{\phi}_1(t) \rightarrow Y$		$\Delta\hat{\phi}_1(t) \leftarrow Y$	
	F-stat	p-value	F-stat	p-value
Currencies				
Return (Daily)	2.10	0.09*	1.26	0.28
Return (Weekly)	0.56	0.45	0.81	0.36
Return (Monthly)	0.06	0.79	0.02	0.86
Return (Quarterly)	1.05	0.45	4.19	0.03**
Volatility (Daily)	1.05	0.38	0.57	0.72
Volatility (Weekly)	0.54	0.70	1.03	0.38
Volatility (Monthly)	0.00	0.95	3.51	0.06*
Volatility (Quarterly)	1.90	0.19	0.99	0.48
Equities				
Return (Daily)	3.97	0.00***	0.97	0.40
Return (Weekly)	0.85	0.35	10.04	0.00***
Return (Monthly)	1.40	0.25	3.10	0.04**
Return (Quarterly)	0.93	0.48	0.61	0.68
Volatility (Daily)	1.37	0.25	4.07	0.00***
Volatility (Weekly)	1.03	0.37	0.84	0.46
Volatility (Monthly)	0.03	0.85	0.72	0.39
Volatility (Quarterly)	2.64	0.10	1.00	0.48

* statistically significant at 10%

** statistically significant at 5%

*** statistically significant at 1%

Table 3.11: Granger-causality test results for the Asian Financial Crisis. Each row represents a test of causality between either returns or realized volatility of MXAS and changes in the AR as derived from currency or equity data, over various time frequencies. The F-stat and p-value found in each test are given for both directions of causality, with $\Delta\hat{\phi}_1(t) \rightarrow Y$ denoting that changes in ϕ_1 Granger-cause the index variable. Conversely $\Delta\hat{\phi}_1(t) \leftarrow Y$ denotes that the index variable Granger-causes changes ϕ_1.

are again maintained as before.

Regression results are presented in Table 3.12 to Table 3.15. Each table corresponds to results with respect to a specific market index. It has been noted here and in existing literature that return correlations among stocks seemed to had been increasing before the Financial Crisis, leading to a conjecture that a measure like the AR could have given an early warning signal. In our experiment we are unable to find a statistically significant relationship between changes in the correlation of components of the NYK and one-period-ahead return or volatility of the SP500 index (see Table 3.12). This is also true for correlations of CDS spreads of major financial institutions. Instead, we find that $\Delta\phi_1$ in equities is coincidentally positively related to volatility. Likewise, positive coefficients in CDS data are found to be statistically significant when changes in AR are contemporaneous or lagging. Returns are negatively associated with $\Delta\phi_1$ at a significant level only when the latter lags by one period. In Section 3.5 a qualitative inspection of co-movement in yields suggested a divergence pattern associated with the crisis. This is confirmed by regression of realized volatility of both BEFINC and EURUSD, as the corresponding coefficients are estimated to be negative. We again observe less consistency in leading AR coefficients than coincident and lagging ones. For example, at a 10% significance level, realized volatility of BEFINC and EURUSD was negatively related to coincident changes in bond yield correlation over three out of four time windows. Finally, the pattern of small p-values for coincident or lagging coefficient in regressing realized volatility is evident in results for the Asian Financial Crisis. Realized volatility is positively related with coincident changes in currencies and equities. Again, we find that the relationship with returns is weak and, in general, the AR does not have predictive power for either index returns or volatility.

3.8 Analysis Caveats

Throughout this chapter we have studied financial data using a framework that involves multiple statistical techniques, including PCA, and inference based on linear regression models or known limiting sampling distributions. These techniques are grounded on assumptions that may not hold in practice. Indeed, financial data is known to exhibit stylized properties that should be considered in applying the methods presented to this point. Thus we proceed with a discussion of caveats to our results and conclusions.

Linearity

It is intuitive that the potential for financial contagion should depend on the strength of co-movement of asset values in the system. However, the use of PCA to measure this as the relative magnitude of the largest eigenvalue assumes that co-movements are linear in nature. That is

Independent Variable	Equities			CDSs		
	Coefficient	t-stat	p-value	Coefficient	t-stat	p-value
SP500 - Leading						
Return (Daily)	0.21	0.99	0.32	-0.11	-0.69	0.48
Return (Weekly)	0.58	1.69	0.09*	0.07	0.32	0.74
Return (Monthly)	0.15	0.44	0.65	-0.34	-0.89	0.37
Return (Quarterly)	0.13	0.33	0.74	-0.47	-0.68	0.49
Volatility (Daily)	0.06	0.49	0.61	-0.03	-0.42	0.66
Volatility (Weekly)	0.12	0.76	0.44	0.15	1.16	0.24
Volatility (Monthly)	0.02	0.12	0.89	0.16	1.10	0.27
Volatility (Quarterly)	0.06	0.40	0.69	-0.06	-0.17	0.86
SP500 - Coincident						
Return (Daily)	0.02	0.13	0.88	0.46	2.78	0.00***
Return (Weekly)	-0.01	-0.03	0.96	-0.41	-1.57	0.11
Return (Monthly)	0.01	0.03	0.96	-0.13	-0.28	0.77
Return (Quarterly)	-0.84	-2.70	0.01**	0.18	0.29	0.77
Volatility (Daily)	0.44	4.27	0.00***	-0.00	-0.01	0.99
Volatility (Weekly)	0.81	4.46	0.00***	0.15	1.09	0.27
Volatility (Monthly)	0.55	2.95	0.00***	0.38	1.30	0.19
Volatility (Quarterly)	0.49	2.59	0.01**	0.61	2.10	0.04**
SP500 - Lagging						
Return (Daily)	0.04	0.20	0.84	-0.31	-2.30	0.02**
Return (Weekly)	-0.95	-3.18	0.00***	0.25	0.86	0.38
Return (Monthly)	-0.94	-2.63	0.00***	0.17	0.65	0.51
Return (Quarterly)	-0.48	-1.69	0.09*	-1.87	-2.70	0.01**
Volatility (Daily)	-0.01	-0.17	0.86	0.17	1.98	0.04**
Volatility (Weekly)	0.25	1.48	0.13	0.19	1.10	0.27
Volatility (Monthly)	0.12	0.70	0.48	0.19	0.76	0.44
Volatility (Quarterly)	0.05	0.46	0.64	-0.20	-0.51	0.61

* statistically significant at 10%

** statistically significant at 5%

*** statistically significant at 1%

Table 3.12: Regression results for the Financial Crisis. Each row represents a t-test that a regression coefficient is significantly different from zero when either returns or realized volatility of SP500 over various time frequencies is regressed on lead and lag observations of $\Delta\phi_1$ as derived from equity or CDS data. In the top section the independent variable leads $\Delta\phi_1$ by one period, in the middle section they are contemporaneous, and in the bottom section the independent variable lags $\Delta\phi_1$ by one period.

Independent Variable	Bonds			CDSs		
	Coefficient	t-stat	p-value	Coefficient	t-stat	p-value
BEFINC - Leading						
Return (Daily)	-0.03	-0.16	0.86	-0.03	-0.13	0.89
Return (Weekly)	0.21	0.45	0.65	0.30	0.79	0.42
Return (Monthly)	-0.86	-2.15	0.03**	0.28	0.57	0.56
Return (Quarterly)	-1.10	-2.35	0.04**	1.26	2.89	0.01**
Volatility (Daily)	-0.25	-0.97	0.32	-0.01	-0.19	0.84
Volatility (Weekly)	-0.51	-2.55	0.01**	0.41	2.63	0.00***
Volatility (Monthly)	-0.21	-0.91	0.36	0.37	1.47	0.14
Volatility (Quarterly)	0.02	0.24	0.81	-0.07	-0.50	0.62
BEFINC - Coincident						
Return (Daily)	-0.03	-0.11	0.90	0.32	1.30	0.19
Return (Weekly)	-0.20	-0.27	0.78	-0.36	-0.87	0.38
Return (Monthly)	1.46	3.94	0.00***	-0.30	-0.50	0.61
Return (Quarterly)	0.80	2.34	0.04**	-0.47	-1.50	0.16
Volatility (Daily)	-0.20	-1.09	0.27	0.02	0.16	0.87
Volatility (Weekly)	-0.38	-1.98	0.04**	0.28	1.38	0.16
Volatility (Monthly)	-0.59	-1.78	0.08*	0.36	1.56	0.12
Volatility (Quarterly)	-0.68	-7.06	0.00***	0.51	3.67	0.00***
BEFINC - Lagging						
Return (Daily)	0.38	1.66	0.09*	0.06	0.25	0.79
Return (Weekly)	0.98	2.06	0.04**	-0.03	-0.09	0.92
Return (Monthly)	0.03	0.08	0.93	-0.65	-1.08	0.28
Return (Quarterly)	2.50	2.89	0.01**	-1.40	-3.73	0.00***
Volatility (Daily)	-0.21	-1.06	0.28	0.31	2.17	0.03**
Volatility (Weekly)	-0.60	-3.01	0.00***	0.39	2.29	0.02**
Volatility (Monthly)	-0.45	-1.16	0.24	0.46	2.31	0.02**
Volatility (Quarterly)	-0.86	-2.59	0.02**	0.35	2.13	0.05*

* statistically significant at 10%

** statistically significant at 5%

*** statistically significant at 1%

Table 3.13: Regression results for the Eurozone Sovereign Debt Crisis. Each row represents a t-test that a regression coefficient is significantly different from zero when either returns or realized volatility of BEFINC over various time frequencies is regressed on lead and lag observations of $\Delta\phi_1$ as derived from bonds or CDS data. In the top section the independent variable leads $\Delta\phi_1$ by one period, in the middle section they are contemporaneous, and in the bottom section the independent variable lags $\Delta\phi_1$ by one period.

Independent Variable	Bonds			CDSs		
	Coefficient	t-stat	p-value	Coefficient	t-stat	p-value
EURUSD - Leading						
Return (Daily)	0.12	1.11	0.26	0.01	0.25	0.79
Return (Weekly)	-0.00	-0.04	0.96	0.23	1.85	0.06*
Return (Monthly)	-0.27	-0.96	0.34	0.32	1.69	0.09*
Return (Quarterly)	-0.72	-7.58	0.00***	0.40	2.43	0.03**
Volatility (Daily)	-0.16	-2.94	0.00***	0.02	0.54	0.58
Volatility (Weekly)	-0.03	-0.41	0.67	0.10	1.62	0.10
Volatility (Monthly)	0.02	0.41	0.68	0.13	2.61	0.01**
Volatility (Quarterly)	-0.02	-0.49	0.62	-0.07	-2.40	0.03**
EURUSD - Coincident						
Return (Daily)	0.02	0.28	0.77	0.01	0.14	0.88
Return (Weekly)	0.29	1.07	0.28	-0.02	-0.11	0.90
Return (Monthly)	0.92	4.11	0.00***	-0.22	-1.19	0.23
Return (Quarterly)	0.72	5.07	0.00***	-0.31	-2.50	0.03**
Volatility (Daily)	0.09	1.89	0.05*	-0.08	-1.11	0.26
Volatility (Weekly)	-0.05	-0.78	0.43	0.02	0.25	0.80
Volatility (Monthly)	-0.14	-1.78	0.08*	0.04	0.91	0.36
Volatility (Quarterly)	-0.10	-2.95	0.01**	0.12	4.67	0.00***
EURUSD - Lagging						
Return (Daily)	0.23	2.60	0.00***	-0.09	-0.87	0.38
Return (Weekly)	0.45	2.51	0.01**	-0.44	-2.96	0.00***
Return (Monthly)	-0.12	-0.57	0.57	-0.37	-1.80	0.07*
Return (Quarterly)	0.58	3.08	0.01**	-0.30	-1.70	0.11
Volatility (Daily)	-0.10	-1.73	0.08*	0.11	1.87	0.06*
Volatility (Weekly)	-0.20	-4.32	0.00***	0.11	1.55	0.12
Volatility (Monthly)	-0.13	-1.44	0.15	0.12	2.86	0.00***
Volatility (Quarterly)	-0.27	-6.11	0.00***	0.10	2.39	0.03**

* statistically significant at 10%

** statistically significant at 5%

*** statistically significant at 1%

Table 3.14: Regression results for the Eurozone Sovereign Debt Crisis. Each row represents a t-test that a regression coefficient is significantly different from zero when either returns or realized volatility of EURUSD over various time frequencies is regressed on lead and lag observations of $\Delta\phi_1$ as derived from bonds or CDS data. In the top section the independent variable leads $\Delta\phi_1$ by one period, in the middle section they are contemporaneous, and in the bottom section the independent variable lags $\Delta\phi_1$ by one period.

Independent Variable	Currencies			Equities		
	Coefficient	t-stat	p-value	Coefficient	t-stat	p-value
MXAS - Leading						
Return (Daily)	-0.20	-1.02	0.30	-0.62	-2.99	0.00***
Return (Weekly)	-0.16	-0.84	0.39	0.33	0.79	0.42
Return (Monthly)	-0.53	-0.91	0.36	-0.15	-0.21	0.82
Return (Quarterly)	0.32	0.44	0.66	0.69	0.41	0.68
Volatility (Daily)	0.02	0.16	0.86	-0.23	-1.60	0.10
Volatility (Weekly)	0.09	0.72	0.46	0.33	1.98	0.04**
Volatility (Monthly)	0.27	1.53	0.12	0.28	1.34	0.18
Volatility (Quarterly)	0.13	0.88	0.39	0.52	1.53	0.14
MXAS - Coincident						
Return (Daily)	0.02	0.13	0.88	-0.51	-2.51	0.01**
Return (Weekly)	-0.04	-0.14	0.88	-0.28	-0.70	0.48
Return (Monthly)	-0.97	-1.90	0.06*	-0.08	-0.12	0.90
Return (Quarterly)	-0.77	-1.07	0.29	-1.69	-1.32	0.20
Volatility (Daily)	0.20	1.66	0.09*	-0.03	-0.21	0.82
Volatility (Weekly)	0.24	1.77	0.07*	0.31	1.65	0.09*
Volatility (Monthly)	0.50	2.06	0.04**	0.48	1.91	0.06*
Volatility (Quarterly)	0.43	2.43	0.02**	0.43	1.16	0.26
MXAS - Lagging						
Return (Daily)	-0.00	-0.01	0.98	0.31	1.56	0.11
Return (Weekly)	-0.76	-2.72	0.00***	0.36	0.84	0.40
Return (Monthly)	-0.20	-0.34	0.72	0.04	0.05	0.95
Return (Quarterly)	-1.25	-1.92	0.07*	-2.24	-2.39	0.02**
Volatility (Daily)	0.20	2.96	0.00***	0.16	1.28	0.19
Volatility (Weekly)	0.16	0.64	0.52	0.16	0.98	0.32
Volatility (Monthly)	0.13	0.99	0.32	0.19	0.52	0.60
Volatility (Quarterly)	0.26	1.36	0.18	0.49	1.31	0.20

* statistically significant at 10%

** statistically significant at 5%

*** statistically significant at 1%

Table 3.15: Regression results for the Asian Financial Crisis. Each row represents a t-test that a regression coefficient is significantly different from zero when either returns or realized volatility of MXAS over various time frequencies is regressed on lead and lag observations of $\Delta\phi_1$ as derived from currency or CDS data. In the top section the independent variable leads $\Delta\phi_1$ by one period, in the middle section they are contemporaneous, and in the bottom section the independent variable lags $\Delta\phi_1$ by one period.

because PCA seeks linearly uncorrelated components. Similarly, statistical inference based on linear regression models, such as in Granger-causality tests or Dimson regression, assumes that the relationship between the variables is linear. In practice financial asset returns and volatility are non-linear, especially during extreme shocks [14]. While our assumptions reflect a simplified market, they allow for a tractable analysis of a high-dimensional problem in finance.

Normality

Financial return distributions exhibit fat tails precisely because of rare, extreme shocks. The assumption that return and realized volatility are normally distributed underestimates the probability of such events. Therefore hypothesis tests in VAR and regression models where this assumptions may lead to conclusions that are not robust.

Stationarity

The AR time series are obtained through PCA on return observations using the covariance method in a rolling-window fashion. That the AR series is dynamic suggests that individual asset data is not covariance stationary. This supports an already well-known stylized fact that financial data exhibits non-stationarity, which impedes the robustness of models fitted to historical samples.

Parameter Estimation

In performing PCA we estimate the correlation matrix as the sample correlation matrix of the latest 252 trading days. Estimating covariances (and correlations) is a challenging task in itself. There is a variety of approaches that impose different degrees of structure on covariance matrices. Whereas ours is completely free of structure, risk-factor based techniques may produce substantially different results. There are also issues with estimating the true eigenvalues of the correlation matrix and comparing them with results from RMT when their assumptions are not satisfied. For example, if the matrix dimensions are too small then properties of estimated eigenvalues will not be compatible with theoretical limits.

Choice of Index

The quantitative analysis of the relationship between correlation and crises in this paper relies on measures of financial distress. To this end we select broad indices that were believed to reflect this information accurately and efficiently. If the chosen proxy for financial distress does not possess these properties then it would serve as poor input into the analysis of Section 3.7.

Chapter 4

Change-Point Analysis

Change-point detection is the process of identifying abrupt temporal changes in a stochastic process [9]. Change is determined in terms of the distributional properties of the underlying process. The problem of detecting abrupt changes in the statistical behaviour of an observed signal or time series is a classical one, whose provenance dates at least to work in the 1930s on the problem of monitoring the quality of manufacturing processes [84]. More recently, this problem has been studied in a wide array of fields including econometrics [3, 4, 5, 12, 16], environmental science [61], finance [10, 18, 73], image analysis [82], medical diagnosis [62] and network security [79, 80], among others. Change-detection serves a broad range of purposes and detection procedures are commensurately diverse. Perhaps at the most basic level, they differ in being *offline* or *online*. Offline algorithms are algorithms which are applied to a sample of historical observations. The purpose of this tool is to detect changes in an *a posteriori* fashion within the sample with the smallest probability of error. On the other hand, online algorithms are designed for real-time monitoring for changes in a stochastic process. Online detection procedures aim to minimize both the false alarm rate and the detection delay, as each new sample is assumed to incur a cost. We will briefly discuss online change-detection here.

Consider a sequence $\{X_t\}$ of i.i.d. real observations and an associated minimal filtration $\{\mathcal{F}_t\}$. Suppose that X_t obeys one of the following two statistical hypotheses:

$$H_0 : \quad X_t \sim P, \quad t = 1, 2 \dots \tag{4.1}$$

$$H_1 : \quad X_t \sim Q, \quad t = 1, 2, \dots \tag{4.2}$$

where P and Q are two distinct distributions with probability density functions p and q, respectively. Let $L(X) = \dfrac{q(X)}{p(X)}$ be the likelihood ratio and define

$$S_t = \sum_{i=1}^{t} \log L(X_i). \tag{4.3}$$

Since $S_t = S_{t-1} + \log L(X_t)$, for $t > 1$ we have

$$\begin{aligned}
\mathbb{E}\left[S_t|\mathcal{F}_{t-1}\right] &= \mathbb{E}\left[S_{t-1} + \log L(X_t)|\mathcal{F}_{t-1}\right] && (4.4)\\
&= S_{t-1} + \mathbb{E}\left[\log L(X_t)|\mathcal{F}_{t-1}\right] && (4.5)\\
&= S_{t-1} + \mathbb{E}\left[\log L(X_t)\right], && (4.6)
\end{aligned}$$

where the second equation holds because S_{t-1} is $\mathcal{F}_{t-1}$-measurable and third equation holds because $\log L(X_t)$ is independent of $\mathcal{F}_{t-1}$. Now, under some probability measure F with density f, we have

$$\begin{aligned}
\mathbb{E}_F\left[\log L(X_t)\right] &= \int_{-\infty}^{\infty} \log\left(\frac{q(x)}{p(x)}\right) f\left(x\right) \mathrm{d}x && (4.7)\\
&= \int_{-\infty}^{\infty} \log\left(\frac{q(x)f(x)}{p(x)f(x)}\right) f\left(x\right) \mathrm{d}x && (4.8)\\
&= \int_{-\infty}^{\infty} \log\left(\frac{f(x)}{p(x)}\right) f\left(x\right) \mathrm{d}x - \int_{-\infty}^{\infty} \log\left(\frac{f(x)}{q(x)}\right) f\left(x\right) \mathrm{d}x. && (4.9)
\end{aligned}$$

Under H_0, we have $F = P$ so that the first term becomes zero and we are left with

$$\begin{aligned}
\mathbb{E}_P\left[\log L(X_t)\right] &= -\int_{-\infty}^{\infty} \log\left(\frac{p(x)}{q(x)}\right) p\left(x\right) \mathrm{d}x && (4.10)\\
&= -D_{KL}(P \parallel Q), && (4.11)
\end{aligned}$$

where $D_{KL}(P \parallel Q)$ denotes the Kullback-Liebler divergence between P and Q. On the other hand, under H_1 we have $F = Q$ and

$$\begin{aligned}
\mathbb{E}_Q\left[\log L(X_t)\right] &= \int_{-\infty}^{\infty} \log\left(\frac{q(x)}{p(x)}\right) q\left(x\right) \mathrm{d}x && (4.12)\\
&= D_{KL}(Q \parallel P). && (4.13)
\end{aligned}$$

Recall that $D_{KL}(A \parallel B) \geq 0$ for any distributions A and B, with strict equality for $A = B$. Thus, in equation (4.6) we have that $\mathbb{E}\left[\log L(X_t)\right] < 0$ under H_0 and $\mathbb{E}\left[\log L(X_t)\right] > 0$ under H_1. It then follows that, under H_1, $\mathbb{E}\left[S_k|\mathcal{F}_l\right] \geq S_l \; \forall \; l \leq k$, so $\{S_t, \mathcal{F}_t\}$ is a submartingale. Similarly, under H_0, $\mathbb{E}\left[-S_k|\mathcal{F}_l\right] \geq -S_l \; \forall \; l \leq k$, so $\{S_t, \mathcal{F}_t\}$ is a supermartingale. In fact, in both cases we have almost sure divergence [67], with

$$S_t \xrightarrow{a.s.} \begin{cases} -\infty & \text{under } H_0 \\ \infty & \text{under } H_1 \end{cases}. \tag{4.14}$$

If we consider the situation that

$$X_t \sim \begin{cases} P & t < k \\ Q & t \geq k \end{cases}, \tag{4.15}$$

i.e. that there is a change-point in the distribution of X_t at time $k < \infty$, then this change will be reflected as a change in the sign of the mean value of the log-likelihood ratio. In the setting described by expression (4.15), a typical objective of an online monitoring scheme is to detect the change while minimizing both the time delay in raising an alarm after the change and the rate of false alarms. Of course there is a tradeoff between the two performance criteria and studies in the area of stochastic control have sought to optimize it. Shiryaev [72] adopted a Bayesian approach whereby the change point is a random variable with a prior distribution. Lorden [52] formulated the problem with a deterministic change point and sought to minimize the worst-case detection delay subject to a lower-bound on the mean time between false alarms. In both cases the authors arrive at optimal solutions. The cumulative sum (CUSUM) procedure [59], which was originally proposed for continuous inspection schemes, was also proved to be optimal for Lorden's formulation by Moustakides [56]. The CUSUM statistic is defined as

$$g_t = \max_{0 \le k \le t} (S_t - S_k) \tag{4.16}$$

$$= S_t - \min_{0 \le k \le t} S_k \tag{4.17}$$

$$= \max \left(g_{t-1} + \log L(X_t), 0 \right), \tag{4.18}$$

where $S_0 \equiv 0$ and S_k is the log-likelihood ratio defined in equation (4.3). The stopping rule is

$$\tau = \inf \left\{ t \ge 1 : S_t - \min_{0 \le j \le n} S_j \ge b \right\}, \tag{4.19}$$

with $b > 0$ representing an alarm threshold. From equation (4.17) it becomes clear that fact that the expectation of the log-likelihood ratio changes signs under different distributions is a key statistical property. It guarantees the almost sure divergence of g_t so that the threshold b will be crossed after the change-point. The effectiveness of the algorithm depends on the size of that change, i.e.

$$D\left(P \parallel Q\right) = \mathbb{E}_Q \left[\log L(X_t)\right] - \mathbb{E}_P \left[\log L(X_t)\right] \tag{4.20}$$

$$= D_{KL}(P \parallel Q) + D_{KL}(Q \parallel P) \tag{4.21}$$

which is noted to also be the *symmetrized Kullback-Liebler divergence* between P and Q. A larger value of D indicates greater dissimilarity between the distributions. That it is symmetric implies $D\left(P \parallel Q\right) = D\left(Q \parallel P\right)$ so the statistic can be thought of as a distance more intuitively than the asymmetric Kullback-Liebler divergence.

There are a couple of challenges that are common in change-point detection procedures. Firstly, information about the distribution of the series – either before the change, after the change or both – is assumed to be known in parametric methods. In many practical cases this is not the case and so assumptions about the underlying distribution need to be made. Secondly, the use of thresholds in the decision function introduces subjectivity and threshold parameters

likely need to be tuned periodically and between different applications. When one or both of the distributions P and Q are unknown, one can resort to nonparametric methods of change-point detection. This approach is often undertaken due to its practicality, even when optimality cannot be established (see [42, 50, 88] for examples). To alleviate robustness issues in dealing with procedures that involve a binary decision on whether change has occurred, we suggest to monitor the symmetrized divergence statistic $D(P \parallel Q)$. As stated before, this statistic represents the degree of dissimilarity between two distribution. Many detection procedures rely on the change of sign of the expected log-likelihood ratio after a change-point. This statistic, which equals the difference between $\mathbb{E}[\log L(X)]$ before and after the change, has therefore been used as a detectability index. Here detectability can be defined in terms of the performance of change-point detection procedures. For example, the average number of samples taken before a decision can be made in some online detection procedures is proportional to the Kullback-Liebler divergence [9]. In the following section we introduce a nonparametric approach for estimating $D(P \parallel Q)$.

4.1 Nonparametric Monitoring Procedure

We adapt a nonparametric estimation method for multivariate stochastic processes in [50]. Let $\boldsymbol{x}_t \in \mathbb{R}^d$ be a d-dimensional time-series and let $\boldsymbol{X}_t = [\boldsymbol{x}_t', \boldsymbol{x}_{t+1}', ..., \boldsymbol{x}_{t+k-1}']' \in \mathbb{R}^{dk}$ be a batch sequence of k observations of $\boldsymbol{x}_t$. Note that $(\cdot)'$ indicates transposition so $\boldsymbol{x}_t$ is a time-varying vector of length dk. In what follows, instead of using a single observation of $\boldsymbol{x}_t$ as a sample instance, $\boldsymbol{X}_t$ is treated as a sample instance in order to capture the temporal correlation that empirical data tends to exhibit. To detect a change in the underlying distribution, a measure of dissimilarity is computed between two groups of temporally-spaced samples as follows. Let $\mathcal{X}_t = \{\boldsymbol{X}_t, \boldsymbol{X}_{t+1}, ..., \boldsymbol{X}_{t+n-1}\}$ and $\mathcal{X}_{t+n} = \{\boldsymbol{X}_{t+n}, \boldsymbol{X}_{t+n}, ..., \boldsymbol{X}_{t+2n-1}\}$ and suppose that $\mathcal{X}_t \sim P$ and $\mathcal{X}_{t+n} \sim Q$. Denote with $D_{\mathrm{KL}}(P \parallel Q)$ the Kullback-Liebler divergence between two distributions P and Q with densities $p(\boldsymbol{X})$ and $q(\boldsymbol{X})$ respectively. Then we are interested in the quantity

$$D(P \parallel Q) = D_{\mathrm{KL}}(P \parallel Q) + D_{\mathrm{KL}}(Q \parallel P) \tag{4.22}$$

with

$$D_{\mathrm{KL}}(P \parallel Q) = \int p(\boldsymbol{X}) \log \left(\frac{p(\boldsymbol{X})}{q(\boldsymbol{X})} \right) \mathrm{d}\boldsymbol{X}. \tag{4.23}$$

Figure 4.1 illustrates the structure of the samples, where the two groups used to estimate divergence are comprised of n samples, each of which is a batch of length k.

The densities $p(\boldsymbol{X})$ and $q(\boldsymbol{X})$ are unknown and must be estimated in a nonparametric fashion. A naive approach would be to estimate each of them separately and then compute the ratio. However, since knowing $p(\boldsymbol{X})$ and $q(\boldsymbol{X})$ implies knowing their ratio but not vice-versa, estimating the ratio directly is an easier task. Therefore, a method proposed in [78] to estimate the density ratio directly is employed in this thesis.

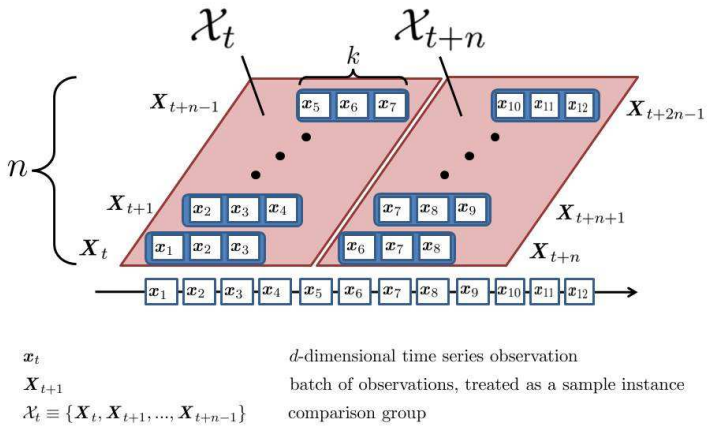

Figure 4.1: Structure of samples used to calculate $D_{\mathrm{KL}}(P \parallel Q)$.

4.1.1 Direct Density-Ratio Estimation

Let us model the density ratio $\dfrac{p(\boldsymbol{X})}{q(\boldsymbol{X})}$ by

$$g(\boldsymbol{X};\boldsymbol{\theta}) = \sum_{i=1}^{n} \theta_i K(\boldsymbol{X}, \boldsymbol{X}_i), \tag{4.24}$$

where $\boldsymbol{\theta} = (\theta_1, ..., \theta_n)'$ is a scaling vector parameter and K is the Gaussian kernel

$$K(\boldsymbol{X}, \boldsymbol{X}') = e^{\frac{\|\boldsymbol{X}-\boldsymbol{X}'\|^2}{2\sigma^2}}, \tag{4.25}$$

where the kernel width, $\sigma > 0$, is to be determined by cross-validation. The parameters $\boldsymbol{\theta}$ are to be learned from the data as those that minimize the Kullback-Liebler divergence between the distributions with densities $p(\boldsymbol{X})$ and $g(\boldsymbol{X};\boldsymbol{\theta})q(\boldsymbol{X})$. That is,

$$\widehat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \int p(\boldsymbol{X}) \log\left(\frac{p(\boldsymbol{X})}{g(\boldsymbol{X};\boldsymbol{\theta})q(\boldsymbol{X})}\right) d\boldsymbol{X} \tag{4.26}$$

$$= \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \left\{ \int p(\boldsymbol{X}) \log\left(\frac{p(\boldsymbol{X})}{q(\boldsymbol{X})}\right) d\boldsymbol{X} - \int p(\boldsymbol{X}) \log\left(g(\boldsymbol{X};\boldsymbol{\theta})\right) d\boldsymbol{X} \right\} \tag{4.27}$$

$$= \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \int p(\boldsymbol{X}) \log\left(g(\boldsymbol{X};\boldsymbol{\theta})\right) d\boldsymbol{X} \tag{4.28}$$

$$= \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \left\{ \frac{1}{n}\sum_{j=1}^{n} \log\left(\sum_{i=1}^{n} \theta_i K(\boldsymbol{X}_j, \boldsymbol{X}_i)\right); \frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{n} \theta_i K(\boldsymbol{X}_j, \boldsymbol{X}_i) = 1, \boldsymbol{\theta} \geq 0 \right\}, \tag{4.29}$$

where in (4.29) the integral is approximated by the empirical estimate and the constraints are added to respect that $g(\boldsymbol{X};\boldsymbol{\theta})q(\boldsymbol{X})$ is a probability density function. Since the problem in (4.29) is convex, there exists a global optimum that can be obtained using the gradient-projection method, among others. Then, the density ratio can be estimated as

$$\widehat{g}(\boldsymbol{X}) = \sum_{i=1}^{n} \widehat{\theta}_i K(\boldsymbol{X}, \boldsymbol{X}_i) \tag{4.30}$$

and an estimate of $D_{\mathrm{KL}}(P \parallel Q)$ is given by

$$\widehat{D}_{\mathrm{KL}}(P \parallel Q) = \frac{1}{n} \sum_{i=1}^{n} \log \widehat{g}(\boldsymbol{X}_i). \tag{4.31}$$

As a simple test of convergence, we estimate D_{KL} for two batches $\boldsymbol{y}_1 = [y_{11}, \ldots, y_{1n}]$ and $\boldsymbol{y}_2 = [y_{21}, \ldots, y_{2n}]$ of n observations simulated independently from the processes $Y_{it} = \theta_i Y_{it-1} + \epsilon_{it}$ with Gaussian white noise. Table 4.1 shows that when $\theta_1 = \theta_2$ the estimate converges zero, its true value, as n grows large. If $\theta_1 \neq \theta_2$ then $D_{KL} > 0$ and $\hat{D}_{KL}$ converges to a positive constant as n grows, albeit at a slow rate (see Table 4.2.)

Convergence test for equal distributions

n	$\hat{D}_{KL}(p(Y_1) \parallel q(Y_2))$
500	0.0286
1000	0.0189
2000	0.0076
4000	0.0051
8000	0.0008

Table 4.1: Estimate of the divergence score between two samples of observations $\boldsymbol{y}_1 = [y_{11}, \ldots, y_{1n}]$ and $\boldsymbol{y}_2 = [y_{21}, \ldots, y_{2n}]$ for different values of n, where $Y_{it} = 0.7Y_{it-1} + \epsilon_{it}$ and ϵ_{it} are independent Gaussian white noise series, for $i = 1, 2$.

Convergence test for unequal distributions

n	$\hat{D}_{KL}(p(Y_1) \parallel q(Y_2))$
500	0.0652
1000	0.0620
2000	0.0597
4000	0.0628
8000	0.0605

Table 4.2: Estimate of the divergence score between two samples of observations $\boldsymbol{y}_1 = [y_{11}, \ldots, y_{1n}]$ and $\boldsymbol{y}_2 = [y_{21}, \ldots, y_{2n}]$ for different values of n. The processes are $Y_{it} = \theta_i Y_{it-1} + \epsilon_{it}$, where $\theta_1 = 0.7, \theta_2 = 0.3$ and ϵ_{it} are independent Gaussian white noise series, for $i = 1, 2$.

This method is completely nonparametric and has the advantage of estimating the density ratio directly, and not the pre- and post-change densities independently. However, its slow rate of convergence may introduce some bias in results given the finite sample size in practice. The kernel width and scaling vector are chosen systematically whereas the parameters which determine the sliding window shape, n and k, are not. These latter two parameters control the size of the rolling windows $\mathcal{X}_t$ and $\mathcal{X}_{t+n}$ and thus act as smoothing parameters as well. The estimated statistic $D(P \parallel Q)$ is to be interpreted as a measure of the distance between P and Q.

4.2 Changes in Correlation Structure

In this section we will relate the concept of statistical change-point detection to the structure of correlations in asset returns. In previous chapters we described the structure of correlations in terms of the strength of co-movement as captured by the relative magnitude of the eigenvalues of the sample correlation matrix. As discussed in Chapter 2, the eigenvalue density of certain models of random matrices tends to a non-random density in the limit that the matrix dimensions grow to infinity. If these densities are known exactly then parametric change-detection techniques can be applied in an online fashion such that a trade-off between the detection delay and false alarm rate is optimized. In practice this is most often not the case, but assuming that a limiting density exists we can attempt to estimate points of high divergence in an online fashion using the nonparametric procedure in the previous section.

Let us revisit the absorption ratio (AR) time series $\{\phi_1(t)\}$ obtained in Chapter 3. For purely random, independently generated Wishart matrices we would expect that observed instances of the first (normalized) eigenvalue would form a stationary series. In fact we would expect this to be true in the limit as the matrix dimensions grow to infinity (with a constant aspect ratio) in all cases where a limiting spectral density exists. In testing the level of the AR for stationarity we find repeatedly for different assets and markets that the series is highly non-stationary, even over relatively short periods of time. This is entirely consistent with the notion that financial crises are associated with a breaking correlation structure. However, non-stationarity in the level of the AR may lead change-detection procedures to produce noisy results. On the other hand, the series $\Delta\phi_1$ have more desirable properties in the sense that non-stationarity due to noise becomes less influential. Increasing divergence between successive observations of the differenced series also maintains the intuitive interpretation of emerging trends in the AR. We shall use the procedure in Section 4.1 to investigate abrupt changes in correlation structure associated with financial crises.

We compute $\Delta\phi_1$ for each of the asset groups studied to this point. We use values of $n = k = 30$ in computing divergence for all cases and plot $\Delta\phi_1$ and the divergence score $D(P \parallel Q)$. We also overlay a plot of the relevant index to compare the divergence score with market conditions. The divergence score increased sharply for financial sector equities and CDS spreads in 2007 (see Figure 4.2 and Figure 4.3). These spikes follow larger, positive values of $\Delta\phi_1$, representing

strengthening of correlations which occurred well before the S&P 500 index reached its crisis lows.

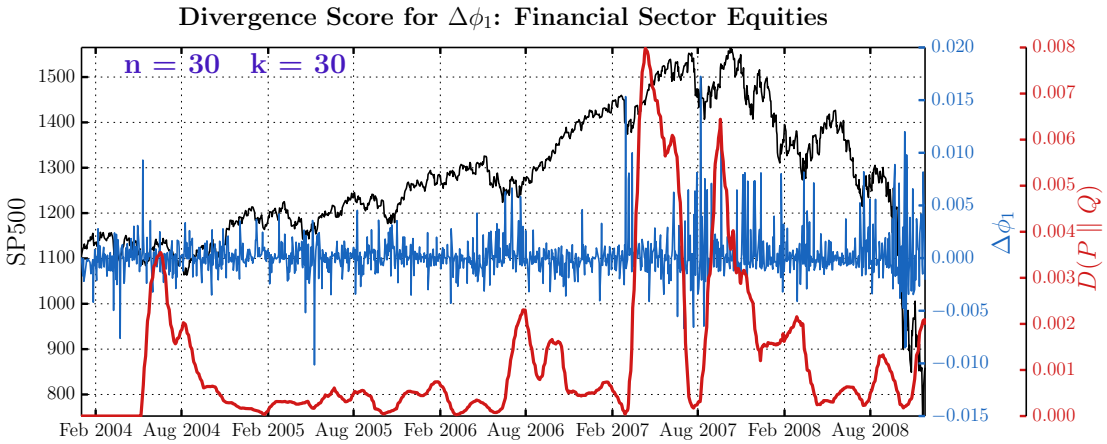

Figure 4.2: Divergence score for the absorption ratio series of the first principal component in NYK constituent stocks.

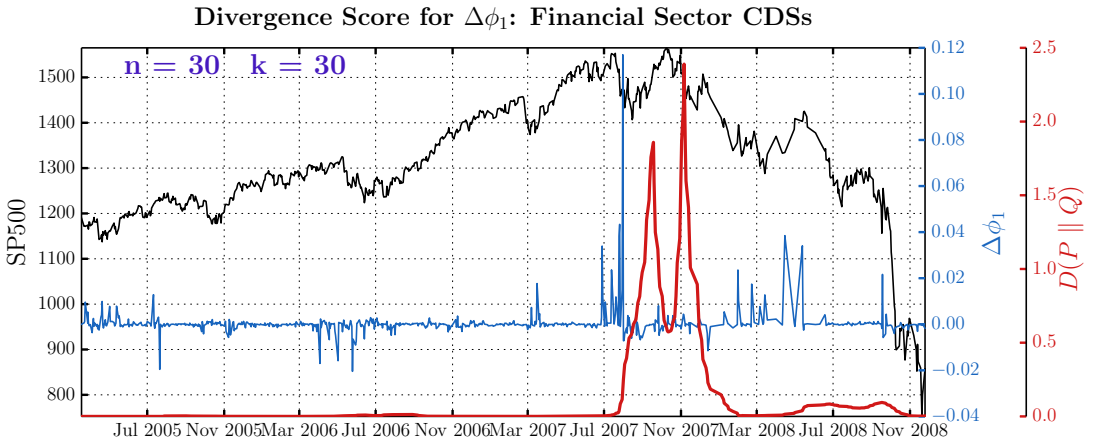

Figure 4.3: Divergence score for the absorption ratio series of the first principal component in bank debt CDS spreads.

During the Eurozone Sovereign Debt crisis the divergence score increased sharply in response to downward shifts in the AR. We find that these shifts are not reflected in the divergence score in bond yield data until the crisis has reached a very developed state (see Figure 4.4). With CDS spreads, based on our data, there is greater potential in detecting a changing correlation structure in an online manner before large drawdowns. This can be observed in Figure 4.5, where a large drawdown in EURUSD starting in the summer of 2011 is preceded by a spike in the divergence score.

Finally, estimating divergence between successive batches of observation of $\Delta\phi_1$ in Asian currencies, we find that divergence increased notably before the start of the Asian Financial Crisis

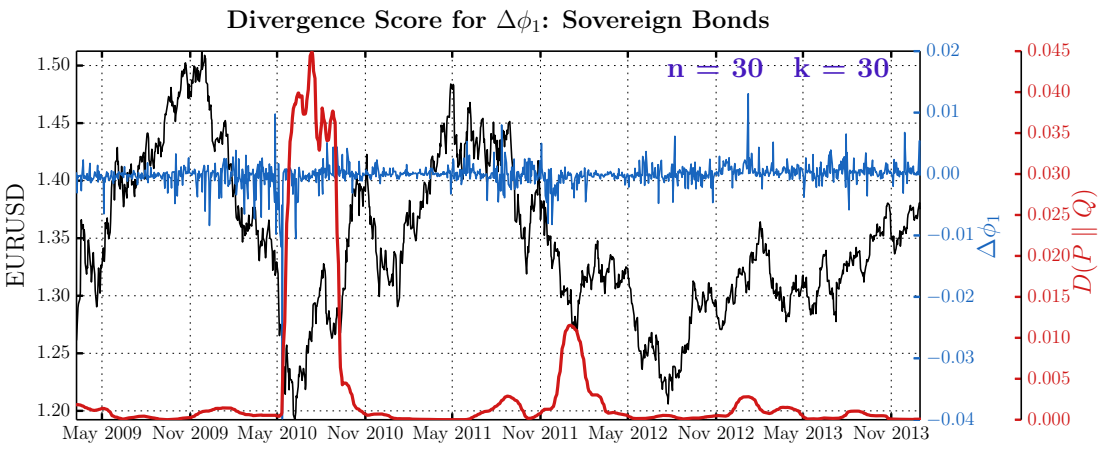

Figure 4.4: Divergence score for the absorption ratio series of the first principal component in European sovereign bond yields.

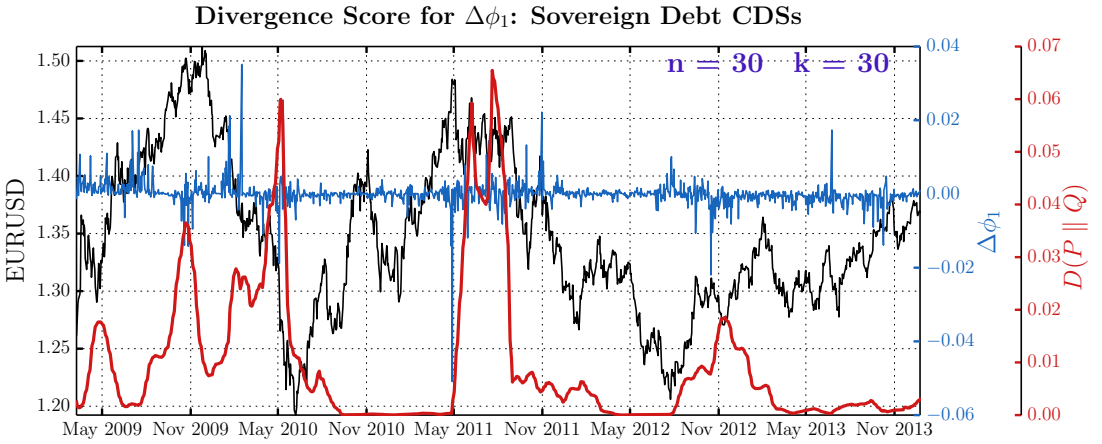

Figure 4.5: Divergence score for the absorption ratio series of the first principal component in European sovereign debt CDS spreads.

(see Figure 4.6). This can also be seen in Figure 4.7, where the results for equity data in this crisis episode are similar, except that the divergence score is noisier. Overall there is an expected lag between observable persistent directional and magnitude changes in $\Delta\phi_1$ and the divergence score because it is estimated from historical observations. We find evidence that despite this lag, a likelihood-based test for a change in the eigenvalue distribution of the correlation matrix of a fixed asset basket between different time intervals may provide early warning about crises. However, financial data is inherently noisy and this results in some robustness issues. For example, results may vary depending on the sampling parameters n and k.

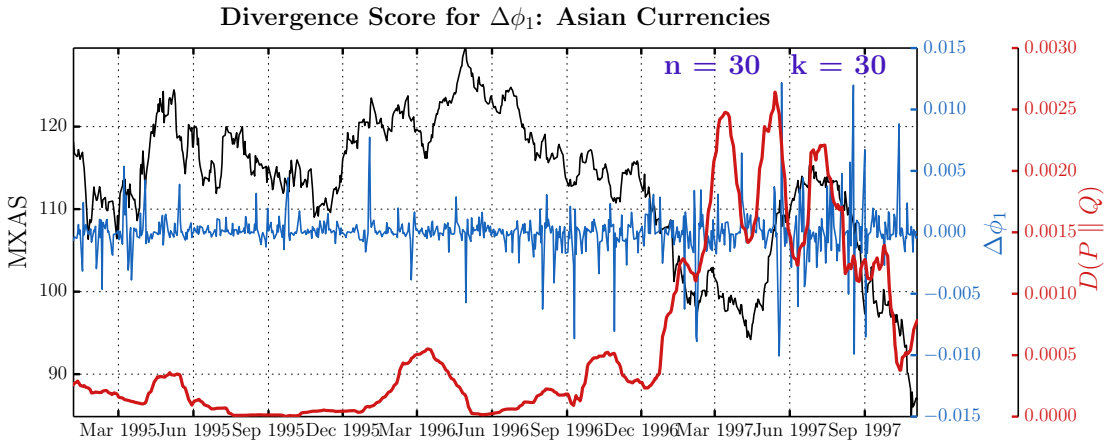

Figure 4.6: Divergence score for the absorption ratio series of the first principal component in Asian currencies.

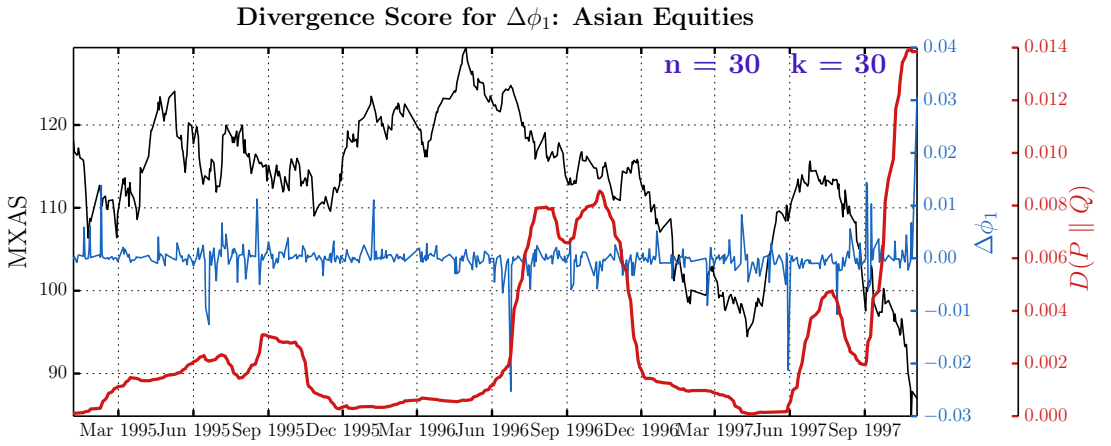

Figure 4.7: Divergence score for the absorption ratio series of the first principal component in Asian equities.

Conclusion

Prior existing studies point to increasing strength of co-movement in returns on assets such as stocks and property before the Financial Crisis. Using the proportion of variance explained by the first principal component in a basket of assets – the Absorption Ratio (AR) – as a measure of correlation, they assert that large drawdowns are associated with increasing correlations, sometimes in advance of financial turmoil, and that the state of higher correlations facilitates contagion.

In this thesis we confirm the findings in the literature about the Financial Crisis. However, in applying a similar framework to other episodes of crisis that involved asset groups that were previously not considered, we find that, generally, financial distress is not necessarily characterized by increasing correlations, but rather a breakdown of the existing correlation structure that may also result in diverging returns. For example, we find evidence that bond yields of countries affected by the Eurozone Sovereign Crisis had diverged. We use Granger-causality tests and linear regression technique to analyze the lead-lag relationship between changes in correlation and financial distress. We find that the AR does not have predictive power with respect to the returns or volatility of a broad index that acts as a proxy for financial conditions. On the contemporaneous or leading basis, evidence on this mounts. In our battery of tests we find that realized volatility is consistently and significantly related to changes in the AR across different time frequencies and asset groups. We statistically confirm that the directional relationship can be both negative and positive.

The AR is found to be richly dynamic and much greater than what would be expected according to random matrix theory if asset returns were random. This suggests the covariance of returns contains structure, or information, and that it is nonstationary. Financial return often have nonlinear relationships with risk factors, and their distributions are often nonstationary and deviate from normality. To the extent that these properties do not obscenely violate our modeling assumptions, the methods here provide some insight into the complicated correlation dynamics of asset returns.

In light of the correlation structure being non-stationary we conduct change-point analysis using a nonparametric technique. By estimating the divergence in the distributions of successive groups of samples through time, we identify time periods that are associated with shifts in the

correlation structure. We find that the divergence score increases either before or coincidentally with systemic financial shocks.

References

[1] Allez, R. and Bouchaud J.-P. (2011). Individual and collective stock dynamics: intraday seasonalities. *New Journal of Physics*, 13, 025010.

[2] Anderson, T. W. (1963). Asymptotic theory for principal component analysis. *The Annals of Mathematical Statistics*, 34(1), 122-148.

[3] Andersson, E., Bock, D. and Frisén, M. (2004). Detection of turning points in business cycles, *Journal of Business Cycle Management and Analysis*, 1(1), 93-108.

[4] Andersson, E., Bock, D. and Frisén, M. (2006). Some statistical aspects of methods for detection of turning points in business cycles, *Journal of Applied Statistics*, 33(3), 257-278.

[5] Andrews, D. W. K., Lee, I. and Ploberger, W. (1996). Optimal changepoint tests for normal linear regression, *Journal of Econometrics*, 70(1), 9-38.

[6] Ang, A. and Chen, J. (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, 63, 443-494.

[7] Bai, Z. and Zhou, W. (2008). Large sample covariance matrices without independence structures in columns. *Statistica Sinica* 18 (2), 425-442.

[8] Bai, Z. D. and Silverstein, J. W. (2010). *Spectral Analysis of Large Dimensional Random Matrices, Second Edition*. Springer, New York.

[9] Basserville, M. and Nikiforov, I. (1993). *Detection of Abrupt Changes: Theory and Applications*. Prentice-Hall, Englewood Cliffs, NJ.

[10] Beibel, M. and Lerche, H. R. (2000). A new look at optimal stopping problems related to mathematical finance, *Statistica Sinica*, 7, 93-108.

[11] Bejan, A. (2005). Largest eigenvalues and sample covariance matrices. M.Sc. dissertation, Department of Statistics, The University of Warwick.

[12] Berkes, I., Gombay, E., Horváth, L. and Kokoszka, P. (2004). Sequential change-point detection in GARCH(p, q) models, *Econometric Theory*, 20(6), 1140-1167.

[13] M. Bijlsma, M., Klomp, J. and Duineveld, S. (2010). Systemic risk in the financial sector: A review and synthesis, *CPB Netherland Bureau of Economic Policy Analysis Paper 210*.

[14] Billio, M., Getmansky, M., Lo, A. W. and Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors, *Journal of Financial Econometrics*, 104, 535-559.

[15] Bouchaud, J.-P. and Potters, M. (2001). More stylized facts of financial markets: leverage effect and downside correlations. *Physica A*, 299, 60-70.

[16] Broemling, L. D. and Tsurumi, H. (1987). *Econometrics and Structural Change*. Marcel Dekker, New York.

[17] Capuano, C. (2008). The option-iPoD. The probability of default implied by option prices based on entropy. *IMF*, working paper 08/194.

[18] Chen, J. and Gupta, A. K. (1997). Testing and locating variance change-points with application to stock prices, *Journal of the American Statistical Association*, 92, 739-747.

[19] Cizeau, P., Potters, M. and Bouchaud, J.-P. (2001). Correlation structure of extreme stock returns. *Quantitative Finance* 1, 217-222.

[20] Conlon, T., Ruskin, H. J, and Crane, M. (2009). Cross-correlations dynamics in financial time series. *Physica A*, 388, 705-714.

[21] Constantine, A. G. (1963). Some non-central distribution problems in multivariate analysis. *Annals of Mathematical Statistics*, 34, 1270-1285.

[22] Corsetti, G., Pesenti, P. and Roubini, N. (1999). What caused the Asian currency and financial crisis? *Japan and the World Economy 11*, 305-373.

[23] Daniel, K. and Moskowitz, T. (2013). Momentum crashes. Working paper, http://www.kentdaniel.net/papers/unpublished/mom10.pdf.

[24] Davis, R. A., Pfaffel, O. and Stelzer, R. (2014) Limit theory for the largest eigenvalue of sample covariance matrices with heavy-tails. *Stochastic Processes and their Applications*, 124(1), 18-50.

[25] De Brandt, O. and Hartmann, P. (2000). Systemic risk: A survey, *European Central Bank*, Working Paper No. 35.

[26] Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366, part 1), 427-431.

[27] Dimson, E. (1979). Risk measurement when shares are subject to infrequent trading. *Journal of Financial Economics*, 7, 197-226.

[28] Doris, D. (2014). *Modeling Systemic Risk in the Options Market*. Ph.D. Thesis, Department of Mathematics, New York University, New York, NY.

[29] Drożdż, S., Grumer, F., Ruf, F., and Speth, J. (2000). Dynamics of competition between collectivity and noise in the stock market. *Physica A*, 287, 440-449.

[30] Edelman A. and Persson, P.-O. (2005). Numerical methods for eigenvalue distributions of random matrices. Working paper, arXiv:math-ph/0501068.

[31] Edelman, A. and Raj Rao, N. (2005). Random matrix theory. *Acta Numerica*, 1-65.

[32] Franses, P. H. and van Dijk, D. (2000). *Non-Linear Time Series Models in Empirical Finance*. Cambridge University Press, New York, NY.

[33] Furman, J. and Stiglitz, J. E. (1998). Economic crises: evidence and insights from East Asia. *Brookings Papers on Economic Activity No. 2*, 1-135.

[34] Geman, S. (1980). A limit theorem for the norm of random matrices. *Annals of Probability*, 8(2), 252-261.

[35] Gopikrishnan, P., Rosenov, B., Plerou, V. and Stanley, H. E. (2001). Quantifying and interpreting collective behavior in nancial markets. *Physical Review* E, 64, 035106-1-3.

[36] Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3), 424-438.

[37] International Monetary Fund. (2009). *Global Financial Stability Report; Responding to the Financial Crisis and Measuring Systemic Risks*. Washington, D.C.

[38] James, A. T. (1960). The distribution of the latent roots of the covariance matrix. *Annals of Mathematical Statistics*, 32, 874-882.

[39] Jin, B., Wang, C., Miao, B. and Huang M.-N. L. (2009). Limiting spectral distribution of large-dimensional sample covariance matrices generated by VARMA. *Journal of Multivariate Analysis*, 100(9), 2112-2125.

[40] Jobst, A. A. (2013). Multivariate dependence of implied volatilities from equity options as measure of systemic risk. *International Review of Financial Analysis*, 28, 112-129.

[41] Johnstone, I. M. (2001). On the distribution of the largest eigenvalue in principal component analysis. *The Annals of Statistics*, 29(2), 295-327.

[42] Kawahara, Y., Yairi, T. and Machida, K. (2007). Change-point detection in time-series data based on subspace identification, *Proceedings of the 7th IEEE International Conference on Data Mining*, 559-564.

[43] Kritzman, M., Li, Y., Page, S. and Rigobon, R. (2011). Principal components as a measure of systemic risk. *The Journal of Portfolio Management*, 37(4), 112-126.

[44] Kwiatkowski, D., Phillips, P., Schmidt, P. and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series has a unit root? *Journal of Econometrics*, 54, 159-178.

[45] Laloux, L., Cizeau, P. and Bouchaud J.-P. (1999). Noise Dressing of Financial Correlation Matrices. *Physical Review Letters*, 83(7), 1467-1469.

[46] Laloux, L., Cizeau, P., Potters, M. and Bouchaud, J.-P. (2000). Random matrix theory and financial correlations. *International Journal of Theoretical and Applied Finance*, 3(3), 391-397.

[47] Lequeux, P. and Menon, M. (2010). An eigenvalue approach to risk regimes in currency markets. *Journal of Derivatives & Hedge Funds*, 16, 123-135.

[48] Lewis, M. (2010). *The Big Short: Inside the Doomsday Machine*. W. W. Norton & Company Inc., New York, NY.

[49] Liu, H., Aue, A. and Debashis, P. (2013). On the Marčenko-Pastur law for linear time series. Working paper, arXiv:1310.7270.

[50] Liu, S., Yamada, M., Collier, N. and Sugiyama, M. (2013). Change-point detection in time-series data by relative density-ratio estimation, *Neural Networks*, 43, 72-83.

[51] Longin, F. and Solnik, B. (2001). Extreme correlation of international equity markets. *The Journal of Finance*, 5(2), 649-676.

[52] Lorden, G. (1971). Procedures for reacting to a change in distribution, *Annals of Mathematical Statistics*, 42(6), 1897-1908.

[53] Marčenko, V. A. and Pastur, L. A. (1967). Distribution for some sets of random matrices. *Math. USSR-Sbornik* 1, 457-483.

[54] Meng, H., Xie, W.-J., Jiang, Z.-Q., Podobnik, B., Zhou, W.-X. and Stanley, H. E. (2014). Systemic risk and spatiotemporal dynamics of the US housing market. *Scientific Reports*, 4, 3655.

[55] Meric, I., Kim, S., Kim, J. H. and Meric, G. (2001). Co-movements of U.S., U.K., and Asian stock markets before and after September 11, 2001. *Journal of Money, Investiment and Banking* 3, 47-57.

[56] Moustakides, G. V. (1986). Optimal stopping times for detecting changes in distributions, *Annals of Statistics*, 14(4), 1379-1387.

[57] Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New York.

[58] NYSE Financial Index. (2014). *NYSE Euronex*. http://www.nyse.com/about/listed/nykid.shtml.

[59] Page E. S. (1954). Continuous inspection schemes. *Biometrika*, 41(1-2), 100-115.

[60] Pan, R. K., and Sinha, S. (2007). Collective behavior of stock price movements in an emerging market. *Physical Review E*, 76, 1-9.

[61] Petterson, M. (1998). Monitoring a freshwater fish population: Statistical surveillance of biodiversity. *Environmetrics*, 9(2), 139-150.

[62] Petzold, M., Sonesson, C., Bergman, E. and Kieler, H. (2004). Surveillance in longitudinal models: Detection of intrauterine growth restriction. *Biometrics*, 60(4), 1025-1033.

[63] Phillips, P. and Perron, P. (1988). Time series regression with a unit root. *Biometrika*, 75, 335-346.

[64] Pillai, K. C. S. (1976). Distribution of characteristic roots in multivariate analysis. Part I: Null distributions. *Canadian Journal of Statistics*, 4, 157-184.

[65] Pillai, K. C. S. (1976). Distribution of characteristic roots in multivariate analysis. Part II: Non-null distributions. *Canadian Journal of Statistics*, 5, 1-62.

[66] Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N., Guhr, T. and Stanley, H. E. (2002). Random matrix approach to cross correlations in financial data. *Physical Review E*, 65(6), 066126.

[67] Poor, V. and Hadjiliadis, O. (2009). *Quickest Detection*, Cambridge University Press, New York, NY.

[68] Preisendorfer, R. W. (1988). *Principal component analysis in meteorology and oceanography*. North Holland, Amsterdam.

[69] Pukthuanthong, K. and Roll, R. (2009). Global market integration: An alternative measure and its application. *Journal of Financial Economics*, 94, 214-232.

[70] Reinhart, C. and Rogoff, K. (2011). *This Time Is Different: Eight Centuries of Financial Folly*. Princeton University Press, Princeton, New Jersey.

[71] Fitch cuts Greece's issuer default ratings to 'RD'. (2012, March 9). *Reuters*, http://www.reuters.com/article/2012/03/09/idUSL2E8E97FN20120309.

[72] A. N. Shiryaev. (1978). *Optimal Stopping Rules*. Springer-Werlag, New York.

[73] A. N. Shiryaev. (2002). Quickest detection problems in the technical analysis of financial data. *Mathematical Finance - Bachelier Congress, 2000 (Paris)*. Springer, Berlin. 487-521.

[74] Silverstein, J. W. (1985). The smallest eigenvalue of large dimensional Wishart matrix. *Annals of probability*, 13(4), 1364-1368.

[75] Silverstein, J. W. (1995). Strong convergence of the empirical distribution of eigenvalues of large-dimensional random matrices. Journal of Multivariate Analysis, 55(2), 331-339.

[76] Smith, R., and Sidel, R. (2010, September 27). Banks keep failing, no end in sight. *The Wall Street Journal*, http://online.wsj.com/news/articles/SB10001424052748704760704575516272337762044.

[77] Solnik, B., Boucrelle, C. and Le Fur, Y. (1996). International market correlation and volatility. *Financial Analysts Journal*, 52, 17-34.

[78] Sugiyama, M., Suzuki, T., Nakajima, S., Kashima, H., von Buenau, P. and Kawanabe, M. (2008). Direct density ratio estimation in high-dimensional spaces, *Annals of the Institute of Statistical Mathematics*, 60(4), 699-746.

[79] Tartakovsky, A. G., Rozovskii, B. L., Blazek, R. B. and Kim, H. (2006). A novel approach to detection of intrusions in computer networks via adaptive sequential and batch-sequential change-point detection methods, *IEEE Transactions on Signal Processing*, 54(9), 3372-3382.

[80] Thottan, M. and Ji, C. (2003). Anomaly detection in IP networks. *IEEE Transactions on Signal Processing*, 15(8), 2191-2204.

[81] Tracy, C. A. and Widom, H. (1996). On orthogonal and symplectic matrix ensembles. *Communications in Mathematical Physics*, 177(3), 727-754.

[82] Trivedi, R. and Chandramouli, R. (2005). Secret key estimation in sequential steganography, *IEEE Transactions on Signal Processing*, 53(2), Part 2, 746-757.

[83] Tulino, A. M. and Verd, S. (2004). Random Matrix Theory and Wireless Communications. *Foundations and Trend in Communications and Information Theory*, 1(1), 1-182.

[84] Wetherhill, G. B. and Brown, D. W. (1991). *Statistical Process Control*. Chapman and Hall, London.

[85] White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4), 817-838.

[86] Wigner, E. P. (1955). Characteristic vectors of bordered matrices with infinite dimensions. *Annals of Mathematics*, 62, 548-564.

[87] Wishart, J. (1928). The generalized product moment distribution in samples from a normal multivariate population. *Biometrika*, 20(A), 32-52.

[88] Yamada M., Kimura, A., Naya, F. and Sawada, H. (2013). Change-point detection with feature selection in high-dimensional time-series data. *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, 1827-1833, 2013.

[89] Yao, J.-F. (2012). A note on a Marčenko-Pastur type theorem for time series. *Statistics and Probability Letters*, 82, 20-28.

[90] Zhang, M., Kolkiewicz, A. W., Wirjanto, T. S. and Li, X. (2013). The impacts of financial crisis on sovereign credit risk analysis in Asia and Europe. The Rimini Centre for Economic Analysis Working Papers 2013, 63.